VOLUME 50
No. 5

WHOLE NO. 225
1938

Psychological Monographs

EDITED BY
JOHN F. DASHIELL
UNIVERSITY OF NORTH CAROLINA

PETERSON MEMORIAL NUMBER

Peabody Studies in Psychology

EDITED BY
LYLE H. LANIER
VASSAR COLLEGE

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION
THE OHIO STATE UNIVERSITY, COLUMBUS, OHIO

To the Memory
of

JOSEPH PETERSON

LATE PROFESSOR OF PSYCHOLOGY
GEORGE PEABODY COLLEGE

TEACHER RESEARCH COUNSELLOR FRIEND

Analytical Student of the Learning Process
Historian of Early Work on Individual Differences
Methodological Pioneer in Race Psychology

Past President of the American Psychological Association
Past President of the Southern Society for Philosophy and Psychology
Past Vice-President, Section I, A.A.A.S.

As a Testimonial to a Distinguished Career in the
Service of Psychology in the South
This Volume is Affectionately Dedicated
By his Students

CONTENTS

JOSEPH PETERSON

Several testimonials to the life and work of Dr. Joseph Peterson were published shortly after his death. It seemed appropriate that a selection of excerpts from these appraisals of Dr. Peterson, as a psychologist and as a man, be included in this memorial volume. These evaluations by his colleagues will serve to convey a partial picture of the man whose memory the contributors to the present volume are proud to honor.

The first selection is from an article by Lyle H. Lanier, a student and research assistant of Dr. Peterson's:

"A hundred years hence the historian may be able to write that psychology grew into a mature science during the first half of the twentieth century. He will possibly observe that by 1950 most psychologists had come to substantial agreement with respect to fundamental postulates, descriptive categories, fruitful problems, and useful methods of investigation. If this prophecy be justified, and recent trends in psychology lead one to hope that it is, then the historian will find few men whose work reflects more adequately the spirit and promise of the period than that of Joseph Peterson. His impatience with futile controversies among the 'schools,' his predilection for crucial experiment rather than arm-chair systematization, his persistent attack upon faculties and mechanical association principles, his organismic conception of adaptive process—these attitudes throw into sharp relief both the man and his time. The career of such a man is, therefore, of great interest to psychologists, as much for its symbolic significance as for its intrinsic contribution."

"The details of educational background and academic affiliation are ordinarily of minor importance in an appraisal of the work and influence of a man of science. In the case of Peterson, however, certain items of this nature serve better to reveal his character than almost any other form of description. He was twenty-six years old and had been married a year when he came to the University of Chicago in 1904 to complete baccalaureate work begun in the colleges of his native state, Utah. He had taught in the public schools of Utah and Idaho since 1899, and his college attendance had been restricted mainly to summer sessions. In spite of almost insuperable financial odds, he managed to remain at the University of Chicago continuously until 1907, earning in that period both the baccalaureate and the doctorate degrees. He returned to Utah in 1907 to teach in Brigham Young University, the school of the Mormon church. But his firm scientific convictions brought him into such conflict with denominational tenets that he resigned in 1911, along with other 'liberals' in the faculty. From 1911 to 1915 he taught in the University of Utah, resigning this position in protest against the administration's treatment of another instructor. It was then that a chance malicious remark by a member of the administration faction provoked a decision which influenced Peterson's entire

career, perhaps adversely from many points of view. This individual observed that Peterson could afford to resign in 'indignation,' since he was fairly certain that the University of Chicago would offer him a position. It so happened that he did receive an invitation to return to his alma mater, but he declined it, both because he did not wish to have the motive for his resignation impugned, and because he felt that the offer might have been made mainly out of consideration for his circumstances. So he accepted a professorial lectureship at the University of Minnesota, receiving a permanent appointment as assistant professor the following year. In the spring of 1918 he was made chairman of the Department of Psychology at Minnesota, but resigned that fall to accept a professorship at George Peabody College for Teachers, in Nashville. The climate, the belief that Nashville would grow into a great university center, and prospects for an extensive development of his own laboratory were important considerations which influenced him to come to Peabody College. He remained there until his death, declining offers which came to him from time to time from other institutions. Of these, none pleased him more than the invitations to return to the Utah colleges which earlier had found his ideas so uncongenial." (Lyle H. Lanier, *Psychol. Rev.*, 1936, **43**, 1–2.)

The following account of Dr. Peterson's professional activities was written by Drs. E. S. Robinson and Florence Richardson Robinson, the latter a contemporary of Dr. Peterson at the University of Chicago:

"Dr. Peterson was an active and prominent member of several learned societies and scientific organizations. He was President of the American Psychological Association in 1934 and member of the Council of Directors of that body from 1927 to 1929. He was a member of the Division of Anthropology and Psychology of the National Research Council from 1926 to 1929 and again from 1932 until his death, at which time he was serving on the Board of the National Research Council Fellowships in the Biological Sciences. In 1931–1932 he was on the Southern Regional Committee of the Social Science Research Council. He was also a member of the Society of Experimental Psychologists, the American Academy of Political and Social Science, the Tennessee Academy, and fellow of the American Association for the Advancement of Science. [*Editor's Note:* He was Vice-President of Section I, A.A.A.S., at the time of his death.] Since coming to Peabody in 1918, Dr. Peterson took a keen interest in psychology and education in the South, and took an active part in several surveys of public school and college teaching. In 1921 he was elected president of the Southern Society for Philosophy and Psychology. Shortly after Dr. Peterson's death President Payne of Peabody, in a letter to President Angell of Yale, voiced an estimate of Joseph Peterson's position: 'You recommended him to me eighteen years ago. In doing so, you conferred the greatest kindness upon higher learning in the South that any man has ever conferred in my lifetime.'"

"In addition to his intense labors in teaching and investigation, Dr. Peterson found time to contribute substantially as an editor of psychological publications. He was an associate editor of the *Mental Measurement Monographs* and a coöperating editor of this Journal (*The American Journal of Psychology*). For many years he contributed to the Journal reports of the scientific papers delivered at the annual meeting of the Southern Society for Philosophy and

Psychology. From 1928 he conducted the book review department of the Journal and supplied a large number of reviews from his own pen. In 1934 he was elected to the arduous and exacting post of editor of the *Psychological Monographs,* a position which he occupied at the time of his death."

"Although Dr. Peterson wrote from time to time upon general problems of education, he was essentially a psychologist's psychologist. He came to the subject with exact talents for mathematics and music and his first original contribution, his doctoral dissertation, was a highly technical monograph on combination tones and other related auditory phenomena. There was a strong interest in auditory phenomena at the time Peterson was a graduate student at Chicago and he plunged into one of the most difficult problems in the field. As we follow his productive record from 1908 on we find that there were few major topics of psychological inquiry in which he did not take an active interest. His experimental and theoretical contributions cover a wide range: combination tones and binaural beats; musical harmony; selective factors in learning; intelligence tests; handling men in the army; transfer of training; rational learning; comparative intelligence of Negroes and Whites; summation of stimuli in learning; history of mental tests; illusion and local sign. It is significant that, of the two pieces of work in preparation at his death, one was an experimental study of inverted vision and the other a critical treatment of the intelligence of different races. Throughout a professional career of nearly thirty years, he had an unusual capacity for entering into the new developments in psychology without losing touch with the more traditional concepts." (Edward S. Robinson and Florence R. Robinson, *Amer. Jour. of Psychol.,* 1936, **48**, 175–176.)

Professor Harvey A. Carr, who was an instructor in the University of Chicago during Dr. Peterson's residence there, writes as follows concerning his characteristics as a man and as a scientist:

"Professor Peterson achieved his recognition on the basis of sheer merit. He was not a self-seeker and he never indulged in advertising and self-exploitation. He was an indefatigable worker—perhaps too much so for his own good. He had high intellectual and scientific standards and was never satisfied with poor and shoddy work on the part of himself or others. He was intellectually honest with himself and never indulged in any form of specious self-deception to attain quick and startling results. He approached all questions with a high degree of objectivity and impersonability of attitude. He evinced an uncompromising devotion to scientific ideals irrespective of its effect upon his personal fortune. In his early professional career he fought the fight for intellectual and scientific freedom, and paid the usual price—which he accepted without rancor, bitterness or complaint. He was invariably sincere, kindly, friendly, and modestly unassuming in all social relations. While he was accustomed, in controversial discussions, to express his own beliefs with vigor, candor, and to subject the beliefs of his antagonists to severe intellectual scrutiny, yet he was essentially too considerate, tolerant, and fair-minded to indulge in personalities. Professor Peterson was a genuine man through and through, and his sterling qualities invariably elicited the respect and admiration of all who knew him." (Harvey A. Carr, *Psychol. Bull.,* 1935, **32**, p. 754.)

The late President Bruce R. Payne brought Dr. Peterson to George Peabody College in 1918 and was president of the college throughout the term of Dr. Peterson's service there. His appreciation of Dr. Peterson as a colleague and faculty-member was expressed in the following words:

"But Dr. Peterson will be remembered by those who were on the Peabody campus during the last seventeen years because of his long hours of work and his unfailing loyalty to professional duties.

"One could set his watch in the morning by his regular and early arrival in his laboratories. He could be expected to return home only after dark each day.

"He attended College functions without question because he was connected with the College. Whether it was a chapel service, a committee meeting or a social gathering, his sense of duty sent him to it. We learned to look for him in all places where he was expected to be.

"He exemplified to all of us in the community the meaning of work and coöperative fidelity. For these reasons I shall miss him more and more with the passing of the mornings when he does not arrive and at the meetings when I shall not see his eager upturned face with its enthusiasm and brilliancy." (Bruce R. Payne, *Peabody Reflector and Alumni News,* 1936, **9**, p. 3.)

BIBLIOGRAPHY OF THE WRITINGS OF JOSEPH PETERSON [1]

Arranged by
LYLE H. LANIER

SENSORY AND PERCEPTUAL PROCESSES

Combination Tones and Other Related Auditory Phenomena. *Psychol. Monog.*, 1908, *9*, No. 39. Pp. xiii + 136.

Titchener on Helmholtz's Explanation of Combination Tones. *Psychol. Bull.*, 1909, *6*, 397–399.

Do Combination Tones Originate in the Middle Ear? *Psychol. Bull.*, 1910, *7*, 169–173.

The Place of Stimulation in the Cochlea versus Frequency of Vibration as a Direct Determiner of Pitch. *Psychol. Rev.*, 1913, *20*, 312–322.

Der Ursprung subjectiver Kombinationstöne. *Annalen der Physik*, 1913, Vierte Folge, Band *43*, 815–816.

Nochmals eine Bemerkung über den Ursprung subjectiver Kombinationstöne. *Annalen der Physik*, 1913, Vierte Folge, Band *43*, 1111–1116.

The Origin of Higher Orders of Combination Tones. *Psychol. Rev.*, 1915, *22*, 512–518.

Moore's 'The Genetic Aspects of Consonance and Dissonance.' *Jour. Phil. Psychol. and Scientific Methods*, 1915, *12*, 694–698.

Illusions of Direction Orientation. *Jour. Phil. Psychol. and Scientific Methods*, 1916, *13*, 225–236.

The Nature and Probable Origin of Binaural Beats. *Psychol. Rev.*, 1916, *23*, 333–351.

Some Striking Illusions of Movement of a Single Light on Mountains. *Amer. Jour. Psychol.*, 1917, *28*, 476–485.

[1] This bibliography is based upon a compilation made by Dr. Peterson a few months before his death. Book reviews and abstracts of articles which were later published have been omitted.

A Functional View of Consonance. *Psychol. Rev.,* 1925, *32,* 17–33.

Local Signs as Orientation Tendencies. *Psychol. Rev.,* 1926, *33,* 218–236.

Chapter on sensory processes (except vision) in revision of Watson's *Psychology from the Standpoint of a Behaviorist.* Philadelphia: J. B. Lippincott Co., 1928.

Habituation Effects of the Equally Tempered Musical Scale. (With W. F. Smith) *Proc. Ninth Int. Cong. Psychol.,* 1929, 338–339.

The Range and Modifiability of Consonance in Certain Musical Intervals. (With W. F. Smith) *Amer. Jour. Psychol.,* 1930, *42,* 561–572.

Motor Reaction

Controls of the Eye-Wink Mechanism. (With L. W. Allison) *Jour. Exper. Psychol.,* 1931, *14,* 144–154.

Learning

Completeness of Response as an Explanation Principle in Learning. *Psychol. Rev.,* 1916, *23,* 153–162.

The Effect of Attitude on Immediate and Delayed Reproduction: A Class Experiment. *Jour. Educ. Psychol.,* 1916, *7,* 523–532.

Experiments in Ball-Tossing: The Significance of Learning Curves. *Jour. Exper. Psychol.,* 1917, *2,* 178–224.

Experiments in Rational Learning. *Psychol. Rev.,* 1918, *25,* 443–467.

Thurstone's Measurements of Variability in Learning. *Psychol. Bull.,* 1918, *15,* 452–455.

The Rational Learning Test Applied to Eighty-one College Students. *Jour. Educ. Psychol.,* 1920, *11,* 137–150.

Tentative Norms in the Rational Learning Test. *Jour. Applied Psychol.,* 1920, *4,* 250–257.

The Backward Elimination of Errors in Mental Maze Learning. *Jour. Exper. Psychol.,* 1920, *3,* 257–280.

A Note on Theories of Learning. *Psychol. Bull.,* 1922, *19,* 443–446.

Learning when Frequency and Recency Factors Are Negative. *Jour. Exper. Psychol.,* 1922, *5,* 270–300.

Tentative Norms of a Simplified Rational Learning Test for Children Eight, Nine, and Ten Years of Age. *Virginia Teacher,* 1922, *3,* 9–10.

Discussion: Frank's Suggestion for a Theory of Learning. *Psychol. Rev.,* 1923, *30,* 402–406.

Johnson's Measurement of Rate of Improvement under Practice. *Jour. Educ. Psychol.,* 1924, *15,* 271–275.

The Elimination of Errors in the Process of Learning. *Jour. Tenn. Acad. Sci.,* 1926, *1,* 11–12.

Limits of Learning by Trial and Error. *Jour. Exper. Psychol.,* 1926, *9,* 45–55.

Forced Adjustment vs. Association in Constructive Learning and Thinking. *Amer. Jour. Psychol.* (Washburn Commemorative Volume), 1928, *39,* 264–282.

Effects of Practice on Individual Differences. (With M. C. Barlow) *Twenty-seventh Year Book of the Nat. Soc. for the Study of Education,* 1928, Part 2, 211–230.

Learning when Frequency and Recency Factors Are Negative and Right Responses Are Painful. *Psychol. Bull.,* 1931, *28,* 207–208.

Effects of Visual Exposure on the Rate and Reliability of Stylus-Maze Learning. (With L. W. Allison) *Jour. Gen. Psychol.,* 1930, *4,* 36–48.

Learning in Children. (Pp. 316–376 in *Handbook of Child Psychology,* edited by Carl Murchison.) Worcester: Clark University Press, 1931. (Revised Edition, 1933, pp. 417–481.)

'Insight' as an Explanatory Concept in Psychology. *Psychol. Bull.,* 1932, *29,* 629.

An Experimental Study of Generalization in Disc Transfer Problems. *Psychol. Bull.,* 1933, *30,* 610–611.

Aspects of Learning. (Presidential Address to American Psychological Association) *Psychol. Rev.,* 1935, *42,* 1–27.

Intelligence and Mental Tests

Growth of Intelligence and the Intelligence Quotient. *Jour. Educ. Psychol.,* 1921, *12,* 148–154.

Intelligence and Its Measurement (Part of a Symposium). *Jour. Educ. Psychol.,* 1921, *12,* 198–201.

Intelligence and Learning. *Psychol. Rev.,* 1922, *29,* 366–389.

Intelligence Conceived as a Mechanism. *Psychol. Rev.,* 1924, *31,* 281–287.

Binet's Early Experiments on Intelligence Testing. *Peabody Jour. Educ.,* 1924, *2,* 87–92.

Early Conceptions and Tests of Intelligence. Yonkers: World Book Co., 1925. Pp. xiv + 320.

How Mental Tests Started. *Industrial Psychol.,* 1926, *1,* 569–579.

Chairman's Report on National Intelligence Tests. National Research Council: *Report of Annual Meeting of the Division of Anthropology and Psychology,* 1928, 30–31.

Social Behavior

Instinctive Tendencies in Man. (A Review of the Literature of Four Years.) *Psychol. Bull.,* 1916, *13,* 412–416.

The Functioning of Ideas in Social Groups. *Psychol. Rev.,* 1918, *25,* 214–226.

Imitation and Mental Adjustment. *Jour. Abnor. and Soc. Psychol.,* 1922, *27,* 1–15.

The Measurement of Beliefs in College Students. *Proc. Ninth Int. Cong. Psychol.,* 1929, 335–336.

Educational and Applied Psychology

The Teaching of Spelling. *Utah Educ. Review,* 1915, *8,* 4–8.

The Neglect of the Modern Languages in the Public Schools. *Jour. Proc. and Addresses of the Minn. Educ. Assoc.,* 1917, 126–129.

Survey of the Public School System. (Part dealing with educational measurements) Paducah School Board, Paducah, Ky., 1919, 71–107.

Getting Results in Teaching. *High School Quarterly,* 1919, *7,* 103–109.

Survey of the New Hanover County Schools. (Joint authorship of section on Educational and Mental Measurements.) Board of Education, Wilmington, N. C., 1921.

Method of Interpreting Results in the Cleveland Arithmetic Test. *Jour. Educ. Research,* 1921, *4,* 280–292.

A Practical and Accurate Method of Ranking Scores. *Peabody Jour. Educ.,* 1923, *1,* 137–143.

Discussion of Dr. Van Wagenen's Paper (A Method of Coördinating Results of Mental and Achievement Scales in Surveys of Achievement and Progress). *Jour. Educ. Research,* 1924, *9,* 350–354.

Effective Distribution of Time and Effort in College Courses. *The Intercollegiate World,* 1926, *1,* 41–47.

Tests and Scaling Methods: E. L. Thorndike's Measurement of Handwriting. (Pp. 220–229 of *Methods in Social Science,* edited by Stuart A. Rice. Chicago: University of Chicago Press, 1931.)

A Constant Error in the Prediction of Achievement. *Psychol. Bull.,* 1928, *25,* 475–476.

Some Effects of Environmental Factors and of Unreliability of Measurement on Rate and Variability in Psychological Development. Washington: National Research Council, Division of Anthropology and Psychology, 1931. Pp. 6.

Race Differences

Comparison of White and Negro Children in Multiple Choice Learning. *Psychol. Bull.,* 1922, *19,* 97–98.

The Comparative Abilities of White and Negro Children. *Comp. Psychol. Monog.,* 1923, *1,* No. 5. Pp. 141.

The Use of a Common Unit in the Measurement of Race Differences. *Psychol. Bull.,* 1923, *20,* 424–425.

Comparisons of White and Negro Children in Certain Ingenuity and Speed Tests. (With H. M. Walker and Lyle H. Lanier) *Jour. Comp. Psychol.,* 1925, *5,* 271–283.

Problems and Results of Testing Negro Intelligence. *Report of Conference on Racial Differences,* Social Science Research Council and National Research Council, 1928, 36–41.

Comparison of White and Negro Children in the Rational Learning Test. *Twenty-seventh Year Book of the Nat. Soc. for the Study of Educ.,* Part I, 1928, 333–341.

Methods of Investigating Comparative Abilities in Races. *Annals of the American Academy of Political and Social Science,* 1928, *140,* 178–185.

Studies in the Comparative Abilities of Whites and Negroes. (With Lyle H. Lanier) *Ment. Meas. Monog.,* 1929, No. 5. Pp. vi + 156.

Results of Individual Tests of Twelve-Year-Old Children on St. Helena Island. *Jour. Tenn. Acad. Sci.,* 1930, *5,* 9–10.

Results of Group and of Individual Tests Applied to the Practically Pure-Blood Negro Children on St. Helena Island. (With C. W. Telford) *Jour. Comp. Psychol.,* 1930, *11,* 115–144.

Basic Considerations of Methodology in Race Testing. *Jour. Negro Educ.,* 1934, *3,* 403–410.

Animal Behavior

Tone and Noise Perception in the White Rat. *Jour. Animal Behavior,* 1916, *6,* 327–329.

The Effect of Length of Blind Alleys on Maze Learning: An Experiment on Twenty-Four White Rats, *Behavior Monog.,* 1917, *3,* No. 15. Pp. 53.

Frequency and Recency Factors in Learning by White Rats. *Jour. of Animal Behavior,* 1917, *7,* 338–364.

White and Tolman on the Elimination of Short and Long Blind Alleys. *Jour. Comp. Psychol.,* 1924, *4,* 443–445.

How Do Animals Learn? *Jour. Tenn. Acad. Sci.,* 1934, *9,* 87–92.

Professional Standards

The Teaching of Psychology in the Teacher-Training Institutions of the South. (With Gladys Dunkle) *Psychol. Rev.,* 1926, *33,* 385–396.

Equipment for the Teaching of Psychology. *Psychol. Rev.*, 1927, *34*, 472–486.

Report of the Committee on the Teaching of Psychology. *Psychol. Bull.*, 1927, *24*, 519–520.

The Development of Instruction in Psychology in the South. *Psychol. Bull.*, 1932, *29*, 392.

Training and Equipment of Teachers of Educational Psychology. Ch. 10 in *Year Book XX, Nat. Soc. of College Teachers of Education*. Chicago: University of Chicago Press, 1932, 101–111.

MISCELLANEOUS

The Psychology of Handling Men in the Army. (With Lt. Q. J. David) Minneapolis: Perine Book Co., 1918. Pp. 146.

How Shall We Find Truth? *Peabody Reflector and Alumni News*, 1930, *3*, 15–16.

Organized and Remotely Directed Service to Man. *Peabody Reflector and Alumni News*, 1931, *4*, 1–5.

DOES PRACTICE WITH INVERTING LENSES MAKE VISION NORMAL? [1]

JOSEPH PETERSON *and* J. KIMBARK PETERSON

As early as 1896, Stratton (**8**) reported an experiment which seemed to prove that the normal inversion of the images upon the retina is not a necessary condition for upright vision. He had worn continuously a system of lenses which had the effect of rotating the whole visual field 180° about the line joining its center to the eye. Rays reaching the eye from a point A not in the center of the field had the same stimulus effect that rays from B, symmetric with A in the fixed line, formerly had. Consequently, an object at A was seen at B and *vice versa.* In this first experiment, the lenses were worn for three days. After an interval of about five months, the experiment was repeated (**10**). This time the lenses were worn continuously for eight days and more detailed records of the reactions were obtained.

Since visual localization can no longer be regarded as an 'innate' capacity, it is quite natural to suppose that vision may become normal again after a long period of seeing exclusively

[1] Indebtedness is gratefully acknowledged to Professor P. H. Ewert for the loan of his inverting lenses, for many months; to Dr. W. H. Boyer and Mr. Aubrey W. Bickley for the information concerning the coördination tests.

Editor's Note: The experiment described in this article was performed by Dr. Peterson during the year which preceded his death. He gave brief reports upon the work at the meetings of the Society of Experimental Psychologists (Yale University, April 4–5, 1935) and of the Southern Society for Philosophy and Psychology (Nashville, April 19, 1935), but he never wrote a complete account of the entire experiment. (A brief news note concerning it appeared in *Science News Letter,* June 8, 1935.) It seemed eminently fitting that a full report of the study be included in this memorial monograph, and, at the request of the editor, Dr. Peterson's oldest son, J. Kimbark Peterson, has written this article. He had assisted Dr. Peterson with the recording of observations during the course of the experiment, and was, therefore, familiar with all of the details of the work.

Insofar as possible Dr. Peterson's diary notes have been reproduced in detail, and the passages in the article which are printed in small type are in his own language. The remainder of the article is the work of J. Kimbark Peterson.

through the lenses. For should it not be possible, in time, to modify the whole system of local signs of vision in such a way that each visual stimulus-complex becomes referred to the place A symmetric with the place B to which it is ordinarily referred? An object normally seen at A would then be seen at A through the lenses, rather than at B. If vision through the lenses is to become normal in this sense, such modification of the 'local signs' is also necessary.

Peterson (5) was the first to propose that 'local signs' are simply orientation tendencies, and to show how they become established. He pointed out that this view places the study of local sign and perception directly in the field of learning, perceptions being, apparently, "reactions which change with new demands of the environment on the individual, just as all other habitual reactions do" (6, p. 8). He, particularly, would have been prompted to accept the conclusion that practice makes inverted vision upright. Yet he saw that Stratton's proof of the conclusion was not quite satisfactory, although it had been accepted for over thirty-five years. Furthermore, certain aspects of the learning situation presented in this experiment were seen to be almost unique and to hold promise of deciding between the organismic and certain other theories of learning (6, p. 5).

Let us review the experiment briefly. For simplicity of exposition, we disregard depth factors. Let the circles in Fig. 1 represent the two-dimensional field of vision. If we bear in mind that an object in the field is seen through the lenses, at least at first, at the point Q symmetric in the center with the point P at which it is normally seen, we can anticipate some of the experiences which confront the subject, S, when he puts on the lenses. When he looks at his feet, S sees them as in the second circle because he would see them without the lenses as in the third circle. His feet seem to belong to someone else who stands facing him. If S turns his head to the right, the external scene seems to shift, more rapidly than S's head, from left to right, seen objects passing on ahead out of view to the right (without the lenses they would pass out of view behind, on the left), and new objects, apparently following these in the shifting of

the environment, entering the visual field from behind. A similar flow of objects across the field in the direction of the rotation seems to occur for up and down turns of the head.

Suppose S kicks a ball in the hallway. While he feels his foot kick the ball forward, he sees the intimate stranger of the second circle kick the ball across the field of vision toward S. But to keep the ball in view as it apparently comes toward him, S has only to raise his head gradually. His own feet are then seen to recede from him and disappear. Parts of the floor really more distant stream into view from behind, and then from

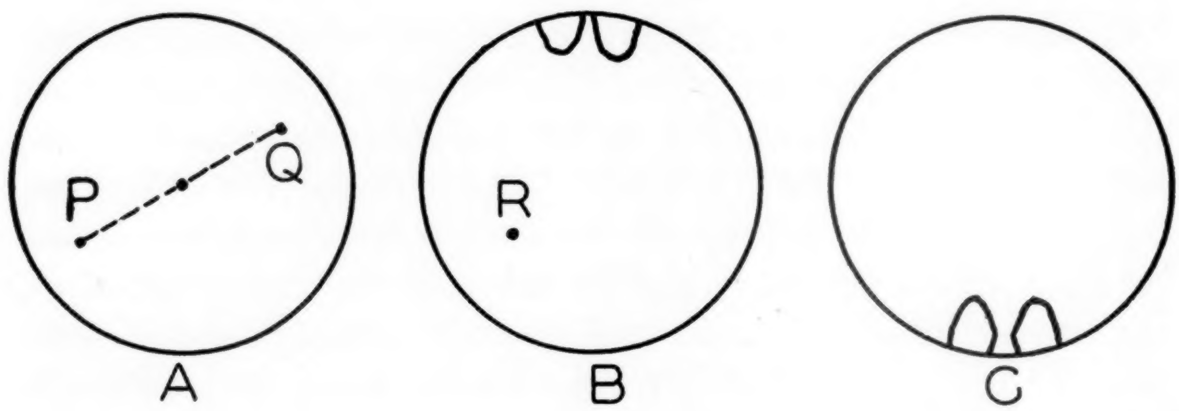


FIGURE 1. Illustrations of the apparent displacement caused by the lenses. (A) Relation between normal position of seen object (P) and the position of the same object as seen with the lenses (Q). (B) Showing the feet as seen through the lenses, in contrast with (C) which shows the feet as seen in normal vision.

underneath, as S continues to raise his eyes. The floor tilts upward and moves into a position overhead. As S raises his head still more, the floor disappears, coming toward him overhead and being replaced by the far wall and then by the ceiling, which remains above him only while he looks sharply upward.

Put into such an environment, S's movements are awkward, uncertain, and ill-adapted. All except the simplest manual operations are tiring, because they require constant checking and guidance. But with several days' practice, S learns to react very well to the new situation. He learns by experience to turn to the left to face something seen on the right. He moves about easily and reaches in the proper direction for objects. He gradually learns to anticipate what he will see when he turns his head

this way or that, or moves thus and so in the same or into another room. In this way he begins to localize objects in a representation-space which is an extension of his visual field—he correlates with the objects of his environment the movements which will bring them, in a predictable way, into the visual field. Objects which he cannot see directly are assigned their place in the system, too. For example, S sees his hand disappear from view in the region of the chest (according to the old localization) and at once feels it touch his scalp. With repetition of this experience, he tends to localize his scalp in this region.

S gradually loses the feeling that the environment shifts too when he moves, and comes to feel that the direction of his movements is such that seen objects drop from the field *behind*, as they do in normal vision. As Stratton found (**10**, p. 481), "the 'absolute' muscular direction cuts no decisive figure in the perception at all." We saw earlier that objects actually leave the moving inverted field at the front edge. Hence the direction of movement is judged (at the stage under consideration) to be just the opposite of the actual direction. Since the new representations place the objects successively seen in the moving field in reverse order with respect to their old positions, no inconsistency traceable to movement is detected.

Before the head and feet were relocalized, S would feel a patent inconsistency when, after looking down at his feet, he raised his head ninety degrees and saw the floor *above* the line of sight. But now, in looking at his feet he feels that he is looking in our up-direction, in the direction in which he sees things fall and in which he has localized his feet. As he raises his head, the floor occupies first the whole of the visual field and later just its upper edge, which is to him the lower. The movement seems to be in our downward direction, in which he now represents the sky. The whole experience is consistent.

For Stratton, the new localizations became spontaneous and immediate. In time, when he looked at his foot, what he saw did not seem to be a strangely displaced image of a real foot somewhere else: it *was* his foot. He could anticipate sensations of contact by watching that foot as he moved; and he could,

later, *feel* them just there where he saw the contact take place. Only when he was in repose and not using his eyes were the old representations dominant. When he was active and seeing, the spontaneous representations of his body were always those of the new system, and if he voluntarily pictured his body in the old way, it seemed unreal and indistinct.

Near the end of the experiment, Stratton's sensations of touch were still somewhat confused. A contact sensation was sometimes felt at both the old and the new place of the part affected. Sometimes a pressure was felt in the new position of the wrong member. For example, a contact with a sheet of paper projecting beyond the pile was distinctly felt in one hand, where it lay in his lap in the field of view, until the stimulating edge was actually seen to be in contact with the other. The feeling then shifted at once to this hand and could not by any effort be brought back to the first one (**10**, p. 360).

At the last, the process of relocalization had progressed so far for Stratton that only the eyes and parts of the head very near them had not been affected. There was a certain degree of harmony between touch and sight, for he usually felt things where he saw them. Movements were made rather freely and confidently. When he was active and had his hands or feet in the field of view, there was no suggestion of any incongruity in the situation. At other times, he felt that something was inverted, perhaps himself, perhaps the environment, but the latter did not seem unnatural. It was like the view a football center gets as he looks back between his legs, ready to snap the ball.

In an article discussing his "upright" vision at length, written in answer to objections by Hyslop, Stratton (**9**) makes clear his expectation that all feeling of inconsistency would have been lost had the experiment been continued much longer. He writes: "But all this is only a transitional state of consciousness. Suppose that the partial reharmonization of my experiences had given place to a complete harmony of tactual and visual perception and to a suppression of my old localizations brought over from the earlier experience—a result toward which the experiments surely point;—I would then feel and see my body unre-

servedly in its new place in the visual field, and in the same relation to the new objects around my body, as existed between my body and surrounding objects in the older experience, viz., my feet on the ground, my head toward the sky, etc." (**9**, p. 186). He considered his vision to be upright in large part—upright in the sense of being harmonious with tactual and motor sensations—and certainly to be so when he was directing movements of his hands or feet in the visual field. If inverted retinal images were necessary for upright vision, it would be impossible to see upright for any period, however short. Hence inverted retinal images are not necessary for upright vision, he reasoned (**8**, p. 616). It was to test the contrary proposition that Stratton performed his first experiment on inverted vision.

It is obviously erroneous to suppose that the 'local signs' of vision had been changed. A period of confusion would certainly attend the changing of such ingrained habits. But at no time during the experiments was there reported any evidence of indistinct, confused, double, or variable visual images. Nor was there any change reported in the appearance of objects from day to day. Even more conclusive proof that the visual local signs were not changed is the fact that when the lenses were finally removed "the visual arrangement was immediately recognized as the old one of pre-experimental days" (**10**, p. 470), although it seemed rather bewildering for several hours because of its contrast with the appearances which had become customary during the experiment.

While most of the text-books lead one to believe that practice with the lenses made Stratton's vision normal (to be distinguished from *natural*), there is, as we have just seen, no evidence from his experience that vision would ever become normal in the experimental situation. At the end of the longer experiment, the older representations seemed still to be the stronger ones, since they were entirely suppressed only when the attention was occupied by a situation in which only the new representations were of practical use. The old representations held sway when the glasses were taken off for the night and before they were put on in the morning. Moreover, when the

lenses were finally removed, no distinct errors in localizing parts of the body occurred, although it caused surprise, more than once, to see the hands enter the visual field from the old lower side (**10**, p. 471). Hence Stratton did not even approximate very closely in his experiment to complete reorientation, with respect to consistency of visual with tactual and motor sensations.

We shall return later to the question whether an extension of the experimental period would probably have made Stratton's vision "upright." We note here that the course of his experiments was just about what would be expected on the basis of a theory which regards learning as the elimination of wrong and the strengthening and organizing of right responses. According to such a theory, there would finally be perfect adaptation to the new environment in a practical way. After such adaptation had been made, the situation would no longer present a visual problem. There could be no question of discarding again the right response habits, of reintroducing confusion, and of changing the local signs of vision. If a subject should go into the experiment with no 'set' against adaptation to the new environment, and if he should eventually, perhaps after passing through Stratton's final stage, develop normal vision (as he well might, according to the organismic theory, to resolve certain inconsistencies and achieve a completeness of response) or show real indication of developing normal vision—that would put right-wrong theories of learning in a difficult situation. It was no doubt with such an idea in mind that Peterson decided to repeat the experiment (**6**, p. 3).

In a study (**3**) of spatially discriminative behavior with inverted vision, on three subjects, one of whom wore inverting lenses continuously for almost 196 hours over a period of fourteen days, Ewert had found no evidence that vision would become normal. Ewert was kind enough to lend to Peterson the instrument used in that study. Unlike Stratton's apparatus, Ewert's (**3**, pp. 197–204) was made for binocular vision. In addition to the right-left and up-down reversals which we have discussed, inverted vision involves a tendency for near-far reversal. Even in monocular vision, alteration of the divergence

of the rays from an object tends to change somewhat its apparent distance, because of the difference in the accommodation of the eye needed to focus the rays. The near-far reversal factor is nowhere mentioned by Stratton, so its effect must have been very slight with his lenses. With the binocular apparatus of Ewert, the near-far inversion is striking. A star appears quite near, " distances are inverted and hollow objects are reacted to as though solid, and *vice versa* " (3, p. 200).

TABLE 1

Date 1934–1935	Time of Putting on Lenses A.M.	Time of Taking off Lenses P.M.	Time Deducted	Time of Wearing Lenses
Dec. 25	9:00	11:00	2 hr. at noon	12 hr.
Dec. 26	7:00	11:30		16 hr. 30 min.
Dec. 27	7:30	11:30		16 hr.
Dec. 28	7:20	11:15		15 hr. 55 min.
Dec. 29	7:10	11:10	2 hr. at noon	14 hr.
Dec. 30	7:00	11:50		16 hr. 50 min.
Dec. 31	7:40	12:00		16 hr. 20 min.
Jan. 1	8:00	11:45		15 hr. 45 min.
Jan. 2	7:20	12:00		16 hr. 40 min.
Jan. 3	7:45	11:05		15 hr. 20 min.
Jan. 4	7:15	11:35		16 hr. 20 min.
Jan. 5	7:20	12:35 (A.M.)		17 hr. 15 min.
Jan. 6	7:40	11:40		16 hr.
Jan. 7	8:00	3:20		7 hr. 20 min.
			Total	212 hr. 15 min.
			Average	15 hr. 9.6 min.

After wearing the lenses for a few preliminary periods of several minutes each, Peterson began, on December 25, 1934, to wear the lenses continuously. Until he removed the glasses on January 7, 1935, he opened his eyes only when wearing the lenses, except for one time, on December 29, in an emergency. On that occasion the whole furnace room seemed to whirl about him, so he put the lenses back on at once. From the impressions recorded in the notes; from the fact that an after-experiment in August showed that the new habits are very persistent; and from the fact that very little was seen, due to the whirling sensation, it seems that this momentary interruption of the experiment was without effect.

Table 1 is the time record for the experiment. On each of two days, time was deducted for a nap at mid-day. Through-

out the whole experiment the subject tried to be active and to avoid thinking analytically about the experiment. In getting from place to place in the later days of the experiment, he took care, by choice of time and route—and even, in some instances, by waiting, with the person accompanying him, in one of the campus side paths behind tall bushes—to avoid as much as possible meeting people, especially those who did not know of his wish not to discuss the experiment.

At times, as will be seen, attempts were made to " force the environment." Attempts at adaptation along this line always conformed to the subject's general attitude. They were made in a spirit of spontaneous reaction to the queer surroundings, with the exclusion of any idea that it was part of an experiment or had any ulterior significance.

It was soon found advisable to dictate rather than to write or type the notes on the experiment, since these activities were so difficult that they were distracting in the attempt to relive the experiences of the day. Although active reasoning about the experiences had been suppressed, reasons for certain appearances occurred spontaneously and have been reproduced in the notes. In certain instances it was seen after the experiment that simpler considerations would suffice for explanation, or that the reason given was false. In reading the notes which follow, it will be observed that the course of the experiment parallels that of Stratton's only at the beginning.

December 25 (Penciled)

After having worn the lenses for short periods (ten minutes or so), I put them on this morning to wear continuously. My preliminary experiences had already acquainted me with what to expect.

December 27 (Dictated to J. K. P.)

This relates to the twenty-fifth, when I started wearing the glasses. The experience was at first very new and interesting. Everything was of course inverted, standing upside down. Left was right, and so on. When I looked down at my feet, the shoes were pointing toward me, and when I tried to kick anything, I missed by a good deal, due to the fact that the foot I saw as my right foot invariably turned out to be the left foot (I had been accustomed to the lateral inversions of the mirror). This of course compelled me in time to move my foot before kicking, and see which was really going to be brought into action. Then by thinking the matter over, I soon learned to gauge the direction somewhat more accurately; but still the getting of skill came very slowly.

One time I stepped up to the side of a table, approaching it on a small rug, a corner of which was near at hand. I saw my feet approaching me on the rug, which was seen somewhat ahead of me when I looked down. That was my first experience with such striking appearances as my walking toward myself. Then looking more directly down, I could see my legs and easily verify that the shoes were really my own, even though of course the illusion of inversion of the feet and the rug persisted.

I was suddenly astonished to see beside me a great drop in the level of the floor. It looked like a drop of two or three feet. At the bottom of that hole were some things which I did not clearly recognize. I had to reach out with my hand before my eyes until the ends of my fingers were directly between me and the objects. Then I gradually extended my hand downward until my fingers came into contact with an ink bottle partly covered by some colored paper. I was surprised to find it really on top of the table beside me, although I had studied the new object somewhat and was not able to interpret it. I had seen some unusually large books lying on the same level, apparently a wooden floor, as that which the ink bottle occupied. By contact I found that these, like the ink bottle, were very near to me, so that my hand touched them soon after I began extending it toward them. To my surprise, the big depression in the floor turned out really to be the table beside me, and the objects were so disproportionately large that I had failed to recognize the partially exposed ink bottle.

While I was already acquainted with the fact that near objects tend to appear far away and distant objects nearer by, it took me some time to overcome all the surprises that these relations afforded. Let me give some examples.

When I went into the hallway leading from my study, I found the floor appearing like some promontory with the walls on both sides projecting downward. And yet I could feel those walls beside me. The end of the hallway seemed to stick out toward me, with the walls receding from it, although those receding walls were really the ones I was touching. It made a rather striking appearance.

When I turned and started to go up the stairway, I saw that the stairs, instead of leading up, led downward, so that I could see each foot, as it moved to a new step, going downward. The vertical part just beyond each step extended downward instead of upward. Yet I noted that I was making progress in the right direction, getting up to the top of the stairway.

I noted that things hanging on the wall in the room in such a way as to cut part of the wall off from vision seemed to be impressed a few inches into the wall, cutting off the wall paper rather irregularly and very close to the edges of the covering objects.

At the breakfast table a number of surprising appearances awaited me. When I looked downward my own dishes and plate seemed to be right side up normally, but as I looked out at other things on the table I noticed that other dishes were inverted so as to make a perfectly clear mound, with things which they contained lying right on the top of the mound. And I noticed that my son's spoon just went to the top of the mound, taking off liquid and other food, while nothing ran down the side. It was a queer thing to see how liquid substances would maintain themselves in such a position.

When I reached out (with permission) to touch such a dish and lift it, sliding my hands over to it carefully on the table, everything adjusted itself perfectly, the dish immediately assuming its proper shape. I found that the peculiar illusion of the dish was due to the fact that the near side of it was the far side as I saw it, and *vice versa.*

This inversion of distances was verified in a particularly interesting way as I was walking along in my study and came to a small box containing lead type, that had been set into the room. The box had an oval iron loop-handle on its top end. I was very much surprised to see the handle down on the bottom of the box, the bottom of the box being depressed considerably below the floor. Only the two distant sides and the bottom could be seen. I was totally unable to see the box as a space-filling object. It seemed to be a mere shell.

When I sat down to handle the box I found that I could correct the illusion. This process was something as follows. I would reach out my hand nearby ahead of the lenses and bring it into the visual field, as was my custom to locate anything, then extend my hand gradually toward the object. Suddenly my fingers came into contact with the handle of the box, on its top, before I expected to feel it. Then with my hands I felt around on the top edges, one of which had been entirely invisible until I verified its existence by contact. I ran my hands downward around all sides of the box until they came into contact with a black steel band which was tightly fastened around the box at its middle (the box stood on end). Then moving my hand directly over the surface of the box, I could see clearly the outline of the top, and the band in its true position, one corner coming nearer to me instead of being the farthest away, as it seemed before. That at once gave volume to the box and made it look natural; but upon withdrawing my hand from the box and viewing the box from a distance, it suddenly assumed the crushed-in position that it had at first—a mere frame. The illusion was there again. I repeated the contact experience with the same correcting effects. Again the illusion suddenly came when I withdrew. It reminded me very much of the reversible stairway illusion in the text-books.

One difficulty was always found in the fact that on looking down perpendicularly onto my desk I would of course see the object fixated in the right position, while as I gradually changed my view to things farther away they would appear inverted and nearer to me than the things I looked directly at below me. The place of break for the inversion was never clearly defined.

At the table this had the effect of making it appear that my own plate or dish, from which I was eating, was beyond the more distant things on the table, so that I had to move my fork or spoon over the several other things, giving me a feeling of embarrassment.

The first day I did not walk out at all, and the strain in my eyes, due particularly to inversion of near-far relations, was considerable. So I took a couple of hours off at mid-day for sleep, which brought much relaxation. The objective world was rigid, however, and its topsy-turvy relationships would not yield. Occasionally in handling things like a spoon, I closed my eyes temporarily and found that things then felt consistent, as usual. In the morning I shaved and washed with my eyes closed, before putting on the lenses.

December 26 (Penciled. Hard to make out. Some lines run into others or go off the paper.)

This morning the strain in my eyes was entirely gone. I shaved with eyes shut. I put the lenses on before that, while I found my clothes, shaving outfit, etc., and again after washing and taking a shower bath, eyes shut.

I ventured outside today to go into the garage to wind the hose around the holder for them. I got to the garage easily by first covering the door with my index finger and keeping the finger on the line with the latch while I walked to the latch. I found the hose easily enough and also the holder, but after some

futile attempts I gave up. At first I started turning the cylinder of the holder toward the hose. . . . I soon discovered the error and corrected it. But the job was impossible because my visual field was too limited to take in enough at a time to untangle the hose, a very long one.

Later in the day K. got me to try to play croquet with him. I failed utterly. I was unable to knock the ball and also to see enough of the field to tell what to shoot at, despite K.'s pointing out the wickets for me.[2] So I had to quit in discouragement. I was cold also.

Inside I walked about a good deal ——ing my vision by means of touch. I spent a good deal of time separating red from black checkers, putting one group on the lid and the other into the box; also in putting checkers into piles one at a time in different positions, for practice in visual-motor ———. . . . tired of this and developed a great dislike, just as one does for car driving at an early stage of practice.

It was pleasant to sit looking out of the window at strange scenes, but I realized that . . . resulted in very little skill if any at all. I practiced at hitting the indoor baseball with the mallet in various positions. . . .

Today there was no eye strain at all, only those other very disagreeable general(?) muscle strains(?).

I made several attempts to write on the typewriter by the touch method, which I had never mastered. Time after time when I felt I had done well my folks laughed at my efforts. See samples.

My human as well as my physical environment would not yield to my efforts to make adjustments. Detailed(?) experiences were very similar to those of the first day.

December 27 (Dictated to J. K. P.)

Today I got up feeling a little ill and was not very active in the forenoon. I looked over the headings in the newspaper a little (I read best holding the paper upside down), but the reversal of distance factors makes the smaller print practically unreadable. The same is true on the typewriter keyboard. If I get my eyes rather near to it, the effect is worse, and when I hold them away some distance, the strength of stimulus is not great enough for clearness. Besides this, my hands throw shadows on the keyboard so that I am compelled to remove them now and then to verify my position. Very frequently, still, I raise the wrong hand when I intend to raise a particular one. Even with my hands held in position on the keyboard, I have considerable difficulty in determining which finger hit the key last, because of the confusion of the hands just noted and also because a partially developed touch method makes my fingers run ahead of my vision of them on the inverted keyboard or even of my attention to them severally. However, I did a little better in writing today than yesterday. See records.

I got the mallet and indoor baseball and immediately on beginning practice knocking the ball, I found that the confusion as to which way to hit was not troublesome today. This seemed due chiefly to the fact that I held the mallet in a position that felt right to my muscular perception for knocking the ball to my left, and let the visual pattern be what it might. I brought the mallet slowly against the ball, and although it seemed to push against the ball from the wrong side, I brought it back and hit the ball in the way that *felt* right, and off it went in the proper direction, to my great satisfaction. I maintained this improvement throughout the day's practice.

[2] The wickets were very hard to see even without the lenses. J. K. P.

My human environment, which had been so perverse yesterday, both regarding my typewriting and my batting of the ball, yielded considerably today, even resulting in praise, especially for the hitting of the ball.

I attempted some forcing of my erroneous perceptions into right forms. The box of lead type to which I referred on the first day still persisted in collapsing, so that the handle on the top of the box dropped down onto the bottom, leaving only part of an empty box with the two distant sides in position. The surrounding black metal band seemed bent back into a right angle fitting tightly against the inner portion of the apparently distant two sides. I grasped the handle of the box with my hand, felt around the top edge, and immediately the box assumed its proper position. Then when my hands were removed from it, it would quickly shift to the illusory position. Again I repeated the contact, getting immediate correction of the illusion, but on looking up a moment (not more than a second), I found that the box was back to the wrong shape again. I repeated these experiences until the box looked normal or immediately assumed the normal position just on my looking at it. I found the cause of the illusion to be that the corner nearest me always appeared at first to be the one farthest away. This tendency I at once overcame by touching that nearest corner and the three nearest edges of the box. Soon I learned to effect by the right fixation the same correcting results without touching the box.

I also learned today not to expect a complete reversal of my visual environment at once. Different parts seem to become more familiar according to my experience at handling them, and it was not so easy for me to determine whether there was really reversal or not in the up-down relations. At the end of my third day, I feel much more confident in my ultimate victory over the illusions, but I have rather grave doubts about gaining this victory in the next four days available.[3]

I noticed that the washbowls and the bathtubs do not seem to poke upward as much as they did earlier, and I have considerably more ease in reaching for and picking up things and in getting around among them.

December 28 (Dictated to J. K. P.)

Today I went out into the garage and wound the hose around the cylindrical holder, which I was unable to do the second day. I had no particular difficulties except that it was hard to get away from one edge of the spool about which the hose was being wound. I got help now and then by closing my eyes and going by the muscular feel of the processes. I also had a walk outside on the street with Wynford after dark. This was mainly for recreation, as it was too dark to see anything clearly but lights.

I spent a good deal of time in the hallway today trying to clear up the illusion of feeling the rail on the right side going down and seeing it on the left, and *vice versa* for going up. By looking down upon the floor where I stood, I saw the floor as below my feet instead of inverted above my head ahead of me. On raising my head so the lenses pointed at a farther and farther distance in the hallway, I was able to prevent the up-down inversion to a considerable extent. The floor holds still for a while and then flops over at the end. I was made extremely tired, not in my eyes, but in my body generally. I am a little discouraged because some things seem harder to do today than yesterday, due to a general fatigue. This may presage an advance tomorrow.

[3] Actually, the experiment was continued eleven more days, not four.

December 29 (Dictated to J. K. P.)

I played in the backyard today with a croquet mallet and ball. I knocked the ball quite successfully and found it readily each time. In the way of general improvements, I saw no change today at all, but due to being shoved around a little by Rhoda, I rather lost my orientation and feel that I went back a little. In an emergency, I suddenly removed the lenses for about half a minute and put them on again. There was evidence of great disturbance—a sort of whirling dizziness—while I was looking with normal eyes. There seemed to be no after effect when the glasses were put on again, but there was a very peculiar feeling in my head while they were off.

December 30 (Dictated to J. K. P.)

Today I played outside with the ball and mallet as before. Rhoda, watching me, thought I did better than formerly. I believe, though, that that was because I did not raise my mallet to aim by (holding it up to cover the object). I kept my orientation rather well and got in without using my mallet to find my way.

I played a game of checkers with myself and beat myself once and was defeated once. Then I played three games with Wynford, losing one to two. In all my work with the checkers, just piling them in different places, and so on, and especially in actually playing checkers, I am much bothered by my hands getting in the way of my vision and casting shadows which also conceal objects.

One thing came especially strongly to my attention today which in the earlier confusions had not come so explicitly to view. That is the fact that since my lenses are convex, objects move forward in the direction that my head turns, just as the print moves along the way of your glasses when you pass them over print. This, I found today, has the effect of narrowing my visual field even more than it is already limited by the small aperture; because when I move in any direction the objects already seen move along in that direction in the visual field, thus preventing the appearance of other objects.

Before today, I had used the method of keeping my visual field fixed until I moved my fingers around so they covered the object I sought, and had then extended my fingers until I could reach the object. Under that plan, I soon learned that an object in the left side of the visual field would be covered when my finger was put on the right side, and similarly for up-down and all oblique directions. Since this facility has advanced somewhat, especially yesterday afternoon and today, I have been inclined to reach directly for the object; but since I became more careless about holding the visual field steady, it became more difficult to locate objects. When I turn my head in any direction the movement of the objects seen is about two or three times, nearer three, the distance that the image of the aperture moves; that is, when I make a head movement, the visible objects stream in that direction much more rapidly, so that objects constantly stream into my visual field from behind and flow out of it ahead. This is an entirely new phenomenon from what a person sees with his normal eyes, when the objects really go in the reverse direction and not so fast.

With respect to up-down reversals of objects, there is much that is yet indefinite. When I look straight down at my feet, I always see them as down from my head, the toes, of course, pointing toward me. There is no up-down inversion in that case. But as I raise my head my feet disappear, moving away from me, and new objects clearly standing on the floor or ground right side up shift forward (from me) until the floor seems to be first in a vertical posi-

tion with the tops of the objects toward me, and then finally in a horizontal position above me with the objects turned upside down. It is therefore hardly right to say that all objects are seen upside down.

Today I took a ride in the automobile, Wynford driving. I noticed, when looking out on the right side, on which I sat, first on the ground beside me and then ahead of the car, that the car was certainly resting downward on the ground, with its top pointing up in the direction that my head pointed. I was able to extend this rightness of position to other cars nearby in those directions. But when I looked out through the opposite window on the other side (I was on the back seat), those objects were so far away that they were clearly upside down, as was also the Parthenon, which we circled around turning constantly to our left. That kind of inversion of distant objects finally bothers one less and less, but I would not say that it disappears.

The same is somewhat true of right and left. In the last two days I have walked up and down the hallway a great deal, trying to force the left-right inversion into conformity with what I feel. On one side the wall is covered with books and on the other with white plaster. Now when the plaster is on the right, for one direction of walking, I turn to the right and face the plastered wall, even though the books had seemed on that side just before my turning. The plaster moves to the right through the more slowly moving field and is in position in front of me when I face it. Likewise when I turn to the left, I can assure myself that the books are there.

But if I disregard the right-left inversions, I get along quite well holding to the railing on my left even though I see it on the right when looking straight ahead. The more of this kind of experience I have, the more I react to right and left in the way described, successfully in a practical way, and the less I am bothered by the inversions. This probably means that I learn to disregard visual objects that are misplaced and to follow touch and kinesthetic patterns more completely.

A thing now becomes clear that bothered me at first in striking the ball with the mallet. Of course the mallet seemed to go contrary to the direction in which I knew I was hitting the ball, and that, too, was troublesome at first. But I soon learned to follow my feeling as to which way I was hitting and to disregard the visual appearance while fixating the ball during my stroke. It was very hard to follow the ball in its course, however, even though I moved my head immediately after the blow in the direction the ball was knocked. I found today that my head had really been moved too rapidly. When I did catch the ball in my visual field it rolled in from behind, whereas I at first expected to overtake it and have it enter from the front of the field. That now is comprehensible from what I said before about objects moving forward and entering the visual field from behind. The half minute or so in which I saw the world directly in an emergency when I pulled off the lenses I can now understand. My dizziness and fluttering sensations in the head were due to the fact that the world did not then move when my head moved,[4] a condition to which I have recently been habituated. When I finally remove the lenses I shall study this phenomenon more fully.

I am no longer bothered with the compression of the box of type of which I spoke earlier.

In this experience I wish to note especially the unity of individual things

[4] He was afterwards doubtful that he had been right in his recollection that things did not seem to move with his head movements, as well as that the absence of apparent movement would cause the sensations. J. K. P.

which one has learned to use for different purposes. Acquaintance by handling means seeing a thing as we feel it.

Normally when a person reads he moves his head and holds the book still. Now when I read following this same habit, the last part of the lines is projected so far to the right that I frequently find my neck nearly twisted off to see it. K. remarked one time that it was turned "nearly 180°." I have noticed recently that all of that can be corrected if I move the book or paper to the left and hold my head still.

I early discovered, before I recognized explicitly the exaggerated movement of all things about me, that my pulse made everything around me beat so that I could read my pulse from the movement of the objects about me.

December 31 (Dictated to J. K. P.)

Today it rained all day and I did not get out at all. I played seven games of checkers today. At the first sitting I beat Wynford 3–1 with one tie; later I beat him 2–1. I found it extremely straining on the eyes because his part of the board was seen as nearer me and more natural to play on for me than my own. It seemed impossible to get hold of the checkers in a way to handle them with ease. My hands would be in the way of the squares onto which I wished to move the checkers or would throw shadows on them, and I seemed very helpless to improve the conditions. Only by a careful study of the situation was it possible to see clearly that my own checkers began on my side of the board. I did better playing against Wynford than I had done for some time, largely, it seemed to both of us, because it was so hard to make the moves that I had more time to think out the best moves.

I found a strong tendency to close one eye, as I do for reading, as a rule, and work monocularly.

I worked a good deal on illusions, finding that when I stand before the mirror the person I see seems on first sight to be geometrically symmetrical to me; that is, as the image faces me in the mirror, its right hand is on my left, whereas normally right is seen in the mirror as on the right. Only by a rather strong reasoning process can I make myself see the image right side up; namely, by looking down at my feet and recognizing that as I come up from there toward the head, both the image and I are standing right side up. Yet when I suddenly see myself in a mirror I am at first impressed with the vision that the image is geometrically my opposite, which would make the head downward.

I have worked a good deal before the washbowls today, but it is impossible not to see the taps inverted so that the wall side is toward me and the right hand tap is on my left, unless I think of myself as standing where the wall is and looking away from the wall toward where I am. Only on the basis of such voluntary thinking can it be said that there is no right-left inversion. But such a conclusion always seems rather contrary to the perceptual facts.

When I look at the clocks there are always clear right-left and up-down inversions and other disturbances due to near-far inversions. I cannot force these inversions to disappear. The same is true when I read. I try to avoid turning the book or paper upside down and thereby to force myself to read the inverted and right-to-left lines. Practice makes the reading easier, but no amount of practice enables me to correct the inverted appearance, and strains from the near-far inversion are always present to weaken the sight of the letters. I usually read with one eye only, since focusing for near work is of course poor with the near-far inversion.

I have now no trouble at all in seeing the type box as a solid, but there is always an unnatural leaning of it away from me.

January 1 (Dictated to J. K. P.)

It was clear though cold today, so Wynford and I played croquet on the gravel of the driveway, knocking a single ball between us. I made a number of very good strokes today and did better than ever before in knocking the ball (straighter, more firmly, and harder). I also played checkers, and was less bothered by the cramping and awkward straining that I spoke of yesterday. Wynford beat me two games in succession.

I studied today in detail a number of things, both on the table and in the rooms. I learned rather well to get into a chair approached from different points of view, although at one time I started to sit down between two chairs, having a hand on an arm of each. The chairs are just alike, except that one is a rocker, the other not. I also noted that in no case was any dish on the table apparently turned over, as I had remarked before. The box of type is always perceived in its full dimensions.

I studied again a good deal the matter of whether things are inverted or not. There is little to add to yesterday's remarks, except that I am getting to be much more at ease with objects about me, and that right-left inversions are troubling me considerably less. I believe for reasons given before, however, that the reversals are not disappearing, but I am becoming more accustomed to them.

After this eighth day of experience, I have decided to assist at registering tomorrow and possibly to teach my classes Thursday and Friday with the lenses on and also to keep them on Saturday or most of Saturday and to study the question of reversals very carefully.

January 2 (No notes. Between approving registration cards, he practiced writing and drawing circles and dotting them.)

January 3 (Dictated to J. K. P.)

I noted a little specific improvement in getting around the Peabody grounds and the laboratory. The "200" on the door, looked at from the inside, reads "500," exactly as it did (through the lenses) before I started the experiment. While people are still inverted, the phenomenon gets less and less attention, so that I don't think much about it. It is sometimes hard to say just what I do see when people ask me.

Respecting left and right, I guess more successfully now than before, because I remember the points on which I was wrong and correct for them. I know now that looking in the mirror I get a geometric inversion of myself, so my right is on the same side in the mirror as my left actually is (rights on opposite sides). And when I confront a person who asks which is left and right, I am to say "right" for the side appearing on my right.

I taught three hours in succession today, and it seemed to weaken me a great deal. The whole visible field seemed to be dimmed a good deal, and I had difficulty in identifying persons whom I know but slightly.

January 4 (No notes.)

January 5 (Dictated to J. K. P.)

Yesterday I went about the campus grounds and buildings without so much guidance as the day before. I asked Wynford and various students offering to

aid me merely to observe and to warn me only in case of danger, letting me be free to make my way by my own visual cues. I used a cane to aim by so as to locate steps up or down and other irregularities in my path that might be dangerous.

I was very much struck, both yesterday and today, in going about outside more than I had been able to do earlier, by the fact that when I was approaching any stairway going upward, the near first step seemed far away from me and the farther steps nearer by. I located the beginning of each stairway either by aiming with the end of the cane between it and my eyes as described before or by holding the cane as I walked along so near the floor or sidewalk that it would come into contact with the first step some time before I got there.

By the way, I have become acquainted with the number of steps in different stairways better in the last few days than I had done in over sixteen years of residence in this section before. For instance, I have learned that the two stairways from the sidewalk to my own house have, in order, four and five steps; that there are only two on the first approach to the Psychology Building, and one (definitely known before) just as I step in at the door.

I have learned in the twelve days of wearing the lenses to get around much more easily in environments and among objects which are familiar to me. But both right and left and up and down inversions still persist, although they are not so noticeable as formerly. At several places on the College grounds where a walk leads off from the one I am on, it invariably appears to go to the left if it actually goes to the right, and *vice versa*. On familiar ground this error of visual perception does not bother me any more, however, because I maintain a general proprioceptive or muscular orientation to turn right or left as the known situation demands.

In attempting to read a newspaper or a letter, I frequently pick it up upside down and read naturally without being aware at all that it is really inverted. For example, yesterday my secretary had written a letter for me which I was to sign. When I had read it over, I put it down, making a movement to sign, without writing on the sheet. The secretary at once made a movement to prevent my signing, saying "You aren't going to sign it that way, are you?" I should have started to sign before he could interrupt me had it not been for previous experience of needing to move the pen along to see whether it was sloped right to get the proper alignment of signature and writing.

Yesterday evening a curious thing showed me how very incomplete my adaptations were on that, the eleventh day. I started to arrange a calendar which has the months and various arrangements of days so placed on cards that they can be made suitable for any month or year. It took me fully an hour to take out the various cards, select those two that I wanted, January and the month beginning on Tuesday, and put them back on the rack of the calendar in proper position. Many serious difficulties arose. Inversions, right and left and up and down, were the least of them. Specifically, putting a card into the rack with four small bracket holders was a very difficult job.[5] The card had to be pressed in tightly against those below it to go behind the two upper holders, one at each end. I found myself always pressing the opposite edge down or slightly slanting the card obliquely so that it would not fit in at the two sides.

---

[5] It is a bit difficult with normal vision, since the cards fill the space rather tightly and must be tilted as they are pressed into position, because of a projecting thermometer just above the card rack. J. K. P.

Then I would put my two thumbs on these two holders, trying to push the cards in with my fingers. But I would see the cards pushed off in the wrong direction, which I had not anticipated. I would turn the whole calendar around so as to correct for this difficulty and find that other difficulties as great arose.

I learned finally to be extremely observant, in making only slight movements, of what results each shift of the card would bring about. When I had finished I anticipated correctly that the calendar had to appear wrong side up on my bookcase, as the calendar leaned against the wall, if it was to be in the right position for a normal person to read. It appeared to me that the cards would fall out when the inversion was made, so I held them tightly with my thumb, which was of course unnecessary, since I was really holding the calendar right side up while putting it in place. I was very much fatigued in my wrists and arms when the job was done. Various other similar evidences against real correction of my inverted perceptions come out constantly.

This evening I attempted to knock the indoor baseball along the hallway with the croquet mallet as I had done somewhat imperfectly a few days ago. By being cautious I was able to make ten perfect strokes, placing the ball properly in the other end of the hall. I stopped with this perfect record.

It has been impossible to eliminate the formerly described inversions of the wash bowl.

From these experiences I thoroughly expect to be properly oriented with my natural eyes when the lenses are removed Monday afternoon.

(There are no notes on the experiences of the following two days.)

On January 7 some eye-hand coördination tests were made. These included the well-known star-tracing, A-cancellation, and Freeman Dot Mirror Tracing tests. A fourth test involved a pattern of large, numbered dots on the blackboard, in an area of probably four square feet. The dots were touched as quickly as possible in a definite order with a walking stick. The pattern of clusters of dots of the Freeman test had also been copied on the blackboard, and the various clusters, according to the number of dots in each, were pointed out when the corresponding numbers were called.

Just after the tests, the lenses were finally removed. There was an immediate and severe nausea, with vertigo and vomiting. When the subject could return from the wash-room, to which he had to flee for a few minutes, wearing the lenses, the tests were repeated upon him, without the glasses. Symptoms of nausea were still evident while the tests were being taken. In spite of this fact, the test scores made after removal of the lenses were better than those made just before. As judged by control

tests made on two other subjects,[6] however, habituation to the lenses had considerably impaired the test performances without lenses.

For his own comfort, my father put the lenses back on again in order to continue with the class after the completion of the tests. He wore the lenses for four more hours in the evening, too.

To test the persistence of the habits acquired, the lenses were worn again on August 28, 1935. The following notes describe the results:

August 28 (Typed)

I wore the inverting glasses again for about fifteen minutes. I found it much easier to get around, about as easy as when I took them off nearly eight months ago. I went readily up the stairway and into different rooms and to the window to see the peculiar sights of the earlier experiences. I was able to anticipate every difficulty and meet it as I had learned to do just before removing the lenses. I tried again to bat the indoor baseball with the mallet, as before. In ten trials I readily made proper adjustments. I made one miss in the ten trials; that is, struck the ball so that instead of hitting the other end it struck the side of the hallway. It got very near the end, the stroke being nearly "perfect." I was able to anticipate each inversion in reaching for and in striking the ball, except that one time I reached to the left when I should have gone to the right. Immediately the error was noticed and corrected. I had no trouble in moving right hand or foot, and left, as desired. Ability at the table was fully as good as at the end of practice in January. After the removal of the lenses I was bothered by a considerable strain and almost dizziness [7] about the eyes and head.

This persisted during the writing of these notes [8] and ten or fifteen minutes after, during part of the lunch hour.

August 28 (Typed)

At about 2:30 I put the lenses on and went out to play with a croquet ball and mallet in the back yard. I found it much easier to get around in the yard and

---

[6] Control tests on Peterson himself were not made before the experiment. While it is possible that some were made later after the effects of wearing the lenses had been overcome, results of such tests have not been found. The actual test scores have therefore not been considered significant enough for reproduction. In such a study as this, objective scores are not of such great value. If vision finally becomes normal, then at this stage, the poorer the coördination, the greater the significant learning, possibly. J. K. P.

[7] Since the lenses had been worn only a few minutes, it is possible that these disagreeable symptoms would have been less noticeable if the subject had been completely over an attack of influenza, from which he had not entirely recovered when he left five days later for the Ann Arbor meeting of the American Psychological Association and his fatal vacation trip. J. K. P.

[8] The writing of these [August] notes took place in each case just after taking off the lenses. J. P.

not to get disoriented. There was, however, a difficulty in getting the ball, when I had knocked it some distance from me. The difficulty was not one of direction but of picking up the ball when I got to it. This time (when other things, too, bothered me less than in January) I was able to analyze a difficulty better than then. I noted that I often kicked the ball, toward which I was walking, away from me. Why? It was clearly because of the fact that as I looked down, the part of the ground which was nearest me actually appeared more distant than that farther away. As I looked downward and toward the place that I could "feel" was the position of my feet, I suddenly saw them coming toward me (toes forward; *i.e.*, toward me). And one time I saw one foot thus approaching "me" actually kick the ball away on what seemed to be ground farther ahead of me; but I found it was behind me. This was then recognized as the same sort of trouble I experienced earlier regarding position of things on the table!

The after-experiment revealed a persistence of the habits of adaptation to the experimental situation that is rather remarkable when considered in connection with the fact that eye-hand coördination was better soon after removing the lenses than before. Just after Stratton removed the glasses in his longer experiment, his coördination was much poorer than before. Responding as he did during the experiment, he would turn the wrong way and bump into things in his very effort to avoid them. At the last Stratton was apparently much more practiced in the new habits than was Peterson. The practical absence of near-far inversion effects and the larger field (about 45° instead of about 34.5°) largely account for that fact. The three days of practice in Stratton's earlier experiment had not sufficed to form such persistent habits, for at the beginning of the longer experiment about five months later, no effect of the first was apparent. There was probably some effect in making learning more rapid.

The most striking difference in the experience of the two subjects was qualitative. The queer world became accepted by Stratton, who built consistency from without, by changing his representation of the position of parts of his body. The harmony was disturbed when he reflected on his own relationship to the environment. In the experiences with the box of type, Peterson early forced modification of some of his perceptions with respect to near-far relationships. The queer appearances that would not yield seemed to be more and more ignored, rather

than accepted, and he began to build consistency outward. By projecting himself strongly into the scene, he extended temporarily the domain of upright objects about him in the episodes of the drive around the Parthenon (December 30), of the mirror (December 31), and of the hallway (December 28), which made him very tired in the body generally. The fact that it was sometimes hard for him to tell, when asked, just what he did see (January 3) indicates a state favorable for such an extension.

Does practice with inverting lenses make vision normal? In all the actual experiments on disoriented vision,[9] this slight evidence from Peterson's is the only real evidence that has been obtained for a positive answer. Yet the experiments do not furnish good ground for a negative conclusion, either.

The chief support for this assertion is the interesting case of Douglas and Lanier (2). An approximately rectangular flap of skin, fat, and muscle, left attached to the cheek at one end, was shifted downward about forty-five degrees to replace four-fifths of the lower lip, which had been destroyed in an accident. Nerve fibers innervating roughly three-fourths of the flap were left intact. Contacts on this area were at first referred to the old position of the flap. A long period of confusion intervened. As late as the one hundred forty-third day, reference was in the main to an area intermediate between the old and new positions. Some false localizations persisted until about the three hundredth day.

Purdy (7) reported a case of somewhat similar nature. The terminal phalanx of a middle finger was removed and skin from its volar surface was folded back over the joint onto the opposite surface of the second phalanx. False localization was invariably made over an observation period of ten years.

The duration of each experiment on inverted vision has evidently been far too short to allow for change in the visual local

[9] Wooster's experiments (11) were made with prismatic glasses which rotated the field about twenty-one degrees and which were worn only twenty minutes each day for a very long period. For Brown (1), whose prisms could be adjusted to give any angle of rotation, only one subject (himself) wore the apparatus continuously. For seven days he wore the goggles with the field, which was about ten degrees in size, turned seventy-five degrees.

signs. To be sure, we should not expect, other things being equal, that it would take as long to change the more essential and more practiced responses to visual stimulation as to change the local signs of touch. But there is a factor, peculiar to the experiments on vision, which operated very effectively against change of the local signs. That is the narrowness of the visual field. In making even moderately rapid movements, it was almost or quite impossible to fixate objects.[10] After the experiment had been completed, Peterson wrote that these difficulties of fixation left " little opportunity for corrective orientation habits to develop. In view of these difficulties it would seem profitable to attempt to view objects with a fixed position of the head while . . . exploratory hand movements were in process. My position of perfect quiescence [was then not] conducive to modification of perceptual habits." [11]

Might practice make inverted vision " upright? " [12] Probably, since such vision would constitute a perfect practical adaptation to the new environment. But as we shall show, it seems that Stratton's upright vision would not be free, as he thought, of all feeling of incongruity, so that if attained, it might be only a transition stage leading later to normal vision. The relocalization of the region of the eyes, which he expected as a finishing detail in a largely completed program, would seem to be a drastic transformation amounting in itself to cancellation of the effect of the lenses.

The relocalization of various parts of the body was to the end of giving a certain reality to visual images. It greatly facili-

[10] Early in the experiment Peterson was struck by this observation. Before discontinuing the experiment in the fourteenth day, he had determined to repeat the corresponding one of Young for audition (**12**; **6**, pp. 2–5). The field is not restricted when one puts on the pseudophone. Moreover, the apparatus can be worn indefinitely without greatly interfering with one's ordinary activities. Due to the near-far inversions, reading was so difficult with the inverting lenses that it would have been impossible for the subject to attend to his editorial duties while wearing them. Besides, it was his unreasoned feeling, just before he ended the experiment, that under the circumstances, no matter how long he wore the lenses his vision would not become normal. J. K. P.

[11] From a first draft of an account of the experiment.

[12] Most text-books which mention Stratton's experiments state that his vision did become upright, without explaining his use of the term. This badly misleads the reader who has not seen Stratton's own account.

tates his visual-motor-tactual coördination, or is perhaps rather an aspect of it, for S to think of objects as really being where he sees them. Unseen objects are then thought of as occupying positions determined in a consistent way by those of seen objects. Before he put on the lenses his visual images had that sort of reality. Since the lenses displace the images by rotating them 180°, we see that the relocalization of an object consists in assigning it a position obtained from its old position by the standard inversion.[13] However, the eye itself is not to be relocalized (turned upside down). Indeed, the new representations, we saw, are to give consistency with what the eye sees in its present orientation.

Let us consider the matter from a slightly different viewpoint. Objects are seen in a certain position relative to the eye (which is not a point) as a frame of reference. To displace or rotate the reference frame in a space amounts, so far as relative position is concerned, to keeping the frame fixed and displacing the other points. Relocalizing the eye too would therefore have the effect of rotating the visual field again, hence of canceling the effect of the lenses.

So long as he had very definite conscious or subconscious memory of the preëxperimental situations, S would probably feel at times, as Stratton did, that his eyes and the scene were not relatively right side up. This feeling would be an aspect of the tendency to invert the eyes, either to conform with the other relocalizations or to make the scene appear in the earlier form.

But enough of theoretical considerations. The experiments on inverted vision have been entirely inconclusive with respect to the question raised in the title to this paper and to related questions. The experiments need to be repeated at greater length with several subjects, using the same or similar apparatus. Very large individual differences have been found, and it should be determined to what extent these were due to the use of such

[13] Rotation of space through 180° about the line from the eye to the center of the visual field. In considering binocular vision, we mean by "the eye" either that point between the eyes from which objects appear to be seen, or the pair of eyes, according to the context.

different instruments. It would be interesting to study the effect of different attitudes and predispositions. Some subjects might be instructed and trained in advance not to accept the environment at all and not to lose sight of the fact that they are just performing an experiment, their object being, in a learn-as-learn-can spirit, to attain normal vision, and as quickly as possible.

Without making any relocalization of objects at all, a subject might learn more quickly to make his way in the new environment by imagining himself inverted and the scene normal. The situation is simplified when the subject looks downward, as at the taps of the washbowl, or at his feet, since inverting himself then amounts simply to edging around the field of view, in imagination, and associating himself with its far edge. It is easy for the reader to imagine himself twisted about to fill the shoes of the second circle of Fig. 1, especially if he tilts his head a bit to one side. It is then quite natural to make the correct response, of kicking slightly outward with the right foot, in order to kick a ball seen at C, apparently near the left foot. Such a scheme, we have seen, actually presented itself during the experiment (December 31), but since the scheme amounted to tricking the environment, it was rejected as violating the attitude that had been adopted.

The main service of the experiment here reported was not to call attention to such more obvious questions, but to inaugurate a new program. Since that of Stratton, experiments on visual disorientation have dealt with rather specialized questions concerning perception. Likewise, only direct learning situations, such as memorizing, maze running, and puzzle solving have been studied as problems in learning. A " typical learning curve " has been recognized, for which several mathematical equations have even been suggested, and some deduced from certain simple hypotheses. Of course deviations from the usual regularity of learning have been recognized, with corresponding plateaux and steep rises in the learning curve. It was Peterson's aim to lift experiments on perception out of a very specialized field of interest and to place them in the forefront of the attack on gen-

eral problems in learning. The experiment on inverted vision was to have been followed by others on learning in situations of complex type, learning for which the curve would exhibit declines and troughs, learning under circumstances in which not simply such factors as frequency and recency would be negative (4), but in which the whole early learning would work disadvantageously.

Bibliography

1. Brown, G. G. Perception of depth with disoriented vision. *Brit. J. Psychol.*, 1928, **19**, 117–146.
2. Douglas, B., and Lanier, L. H. Changes in cutaneous localization in a pedicle flap. *Arch. Neurol. & Psychiat.*, 1934, **32**, 756–762.
3. Ewert, P. H. A study of the effect of inverted retinal stimulation upon spatially coördinated behavior. *Genet. Psychol. Monog.*, 1930, **7**, 177–361.
4. Peterson, Joseph. Learning when frequency and recency factors are negative. *J. Exper. Psychol.*, 1922, **5**, 270–300.
5. Peterson, Joseph. Local signs as orientation tendencies. *Psychol. Rev.*, 1926, **33**, 218–236.
6. Peterson, Joseph. Aspects of learning. *Psychol. Rev.*, 1935, **42**, 1–27.
7. Purdy, D. M. Tactual space perception in translocated tissue. *J. Gen. Psychol.*, 1934, **10**, 227–229.
8. Stratton, G. M. Some preliminary experiments in vision without inversion of the retinal image. *Psychol. Rev.*, 1896, **3**, 611–617.
9. Stratton, G. M. Upright vision and the retinal image. *Psychol. Rev.*, 1897, **4**, 182–187.
10. Stratton, G. M. Vision without inversion of the retinal image. *Psychol. Rev.*, 1897, **4**, 341–360 and 463–481.
11. Wooster, Margaret. Certain factors in the development of a new spatial co-ordination. *Psychol. Monog.*, 1923, **32**, No. 146. Pp. 96.
12. Young, P. T. Auditory localization with acoustical transposition of the ears. *J. Exper. Psychol.*, 1928, **11**, 399–429.

FACTORS AFFECTING SPEED IN SERIAL VERBAL REACTIONS

J. Ridley Stroop
DAVID LIPSCOMB COLLEGE

Except in the case of the pre-school child, experimental results have shown consistently that naming colors requires more time than reading the names of colors. Several conflicting theories have been offered in explanation of this difference. The hypothesis which the present paper proposes to test was developed by the late Dr. Joseph Peterson (6, p. 281), who interpreted the difference "as due to the fact that to the written words 'red,' 'green,' etc., the subjects have as a rule given in the past but the one response of pronouncing (vocally or subvocally) the names of these colors; whereas on seeing the colors red, blue, green, etc., they have responded in many different ways, as grasping and eating, handling, perceiving and admiring, etc. In the case of the words, then, but one specific response-habit has become associated with each word, while in the case of the colors themselves a variety of response tendencies has developed."

During a conversation with the writer, Dr. Peterson, shortly before his death, voiced the wish that some one would test out his explanation. He suggested that it might be done satisfactorily by having one group of subjects repeat many times the *same* response to a stimulus symbol and another group make the same number of *varied* responses to the same stimulus. Since the writer had carried out some experimental work on the question in Dr. Peterson's laboratory, he felt that it would be appropriate to comply with this wish in preparing a paper for this memorial publication.

Psychological literature contains several other explanations of the speed difference in question. Cattell (2) and Lund (5) attribute the difference to 'practice' in reading; while Brown (1)

insists that it does not depend upon practice but upon the difference in the association processes involved. Woodworth and Wells (9) account for the difference by the mutual interference caused by the equal readiness of the five color names; but Garrett and Lemmon (3) hold that the interference is caused by an equal readiness of the color recognitive processes. Stroop (7), by an interference method, compared the strength of the associations between colors and the naming response with that between color names and the reading response, and found the latter form of association to be much stronger. Assuming that these associations are products of training, the results are interpreted as supporting Peterson's explanation. Ligon (4) declares that all former explanations are untenable and explains the difference by a common factor which he designates as "a learned reaction," and two special factors (color naming and word reading) which he describes as "organic in nature." Stroop (8) asserts that Ligon's new hypothesis is based upon misleading statistical operations rather than upon acceptable experimental data.

The results of three experiments are presented in this paper. The first, representing an effort to comply with Dr. Peterson's wish, compares the effects of repeating the *same* response to a stimulus with those obtained by repeating the same number of *different* responses to the stimulus. The second experiment compares speed in naming colors and reading names of colors with speed in naming objects and reading names of objects. The third experiment approaches the problem through the medium of another form of reaction: it compares the reaction times to colors and to color names, in card sorting. Detailed descriptions of the materials used in each of the three experiments are included below in the respective sections.

Experiment I

Problem. What will be the effect upon reaction time if one of two equivalent groups of subjects practices the same response to a stimulus while the second group practices an equal number of heterogeneous responses?

Materials. It was necessary that the symbols used as stimuli be entirely unfamiliar to the subjects. For this reason the symbols shown in Figure 1 were chosen. They were printed from

FIGURE 1. Symbols Used in Experiment 1

Blem Nish Hool Nomy Gesic

type formerly used by printers for making borders or other decorative designs. They are accompanied by their assigned nonsense names. Five symbols were used, to correspond in number to the five colors that have been used in studies of color naming and word reading.

For serial presentation test sheets were prepared on which were printed five lines of ten symbols each, with each symbol appearing twice to the line. They were arranged to avoid regularity of occurrence, and were spaced regularly on 8½ inch lines. Eight serial orders were printed. For discrete presentation, cards 2¼ by 3¼ inches were prepared, with a single symbol on each card. The cards were arranged 50 to the deck.

Subjects and Procedure. Forty sophomore college students in psychology served as subjects. They were separated into two groups of 20 each, equated for sex, intelligence, and combined score on a 'form naming' and a 'reading names of forms' test. The last two tests were used in order to equalize the two groups with respect to speed in calling words and naming symbols. A coin was flipped to determine which group should practice the reading response; this group will be referred to as the experimental group. The other group, which made varied responses to the same stimuli, will be called the control group.

The experimental work was individual. It continued through 7 sittings: 3 sittings per week on alternate days. No work was done on Sunday. The introductory part of the first day's work was the same for the two groups of subjects. The subject was seated at a table, on which, arranged in a single row, were five small cards, each bearing one of the five symbols accompanied by its nonsense name. The subject looked at each symbol care-

fully and called its name aloud. The row was read twice in this manner. Then he was given a deck of cards (described in the materials above) which he dealt upon the table, calling aloud the name of the symbol on each card. Until the subject was able to call the name of the symbol from memory he was permitted to refer to the row of cards on the table. The time for each dealing was taken with a stop watch, in order to keep the subjects conscious of the speed factor.

After completing the introductory exercise as described, the regular practice series was begun at the same sitting. Each member of the experimental group read four sheets of 50 symbols each, under the following conditions. On the signal "Ready! Go!" the subject, who was holding the symbol sheet face down, turned it and began to read aloud. The experimenter followed the reading by use of a similar sheet, and timed each trial with a stop watch. A short rest period was allowed after the reading of each sheet. This procedure was repeated at the five succeeding experimental periods, making a total of 1200 verbal reactions, or 240 verbal responses to each symbol. Six different arrangements of the materials were used.

Each member of the control group also made 1200 reactions, in addition to the introductory work. The work of the control group varied so much that it will be necessary to describe each day's exercise. Each member of the group continued the introductory work of dealing cards and naming the symbols as they appeared, using four decks of cards. Thus 200 verbal responses were made to the symbols presented discretely. At the second sitting the cards were sorted into piles according to symbol. The cards in the respective piles were then arranged to facilitate being picked up singly, and as the experimenter called the names of the symbols in the order shown on one of the test sheets, the subject collected the cards into a deck according to that order of arrangement. Two decks of 50 cards were used; hence, 100 sorting responses and 100 collecting responses were made. At the third sitting the five small cards bearing the symbols and names were placed on the table before the subject at a convenient position so that when the finger tips of the subject's

right hand were placed naturally on the table each rested before one of the five symbols. The cards of a regular deck were dealt to the table by the experimenter. The subject tapped the table with the finger before the symbol on the table corresponding to the one dealt. Four decks of 50 cards each were used totalling 200 tapping reactions. At the fourth practice period the work of the second day was repeated, which consisted of 100 card sorting responses and 100 collecting responses. At the fifth practice period the subject dealt four decks of 50 cards each and responded to each symbol by calling a number. The number assigned to a symbol was determined by the order in which five specimen cards were placed on the table, the order being the same as in the third practice period. For the sixth practice period the subject was given a sheet of 50 symbols and was instructed to check with a pencil the symbol named by the experimenter. He was informed that each symbol appeared twice in each line, and each line was completed before the next one was begun. Four sheets of symbols were used, making a total of 200 reactions.

Summarizing the practice exercises, the control group made the following responses during the respective practice periods: (1) 200 naming responses; (2) 100 card sorting and 100 collecting responses; (3) 200 tapping responses; (4) 100 card sorting and 100 collecting responses; (5) 200 'number naming' responses; and (6) 200 checking responses. Thus the control group made a total of 1200 responses to the five symbols during the practice series, corresponding to the 1200 naming responses made by the experimental group. The reactions of the two groups were similarly distributed over an equal period of time. All were timed by a stop watch, to impress upon the subjects the importance of the speed factor.

At a final sitting (the seventh) the results of the practice series were checked for both groups. Two new forms of the test sheet of symbols were used, and each form was used twice, alternately, making a total of 200 reactions for each subject. The sheets were timed separately, the average time being taken as the final score.

Results. The results are given in Table 1. The mean time on the final test for the subjects who had practiced the naming response to the symbols (experimental group) was 28.5 seconds, and the probable error of the distribution was 3.10. The mean time on the final test for the subjects who had practiced varied responses to the symbols (control group) was 40.5 seconds and the probable error of the distribution 7.42. The difference between the means was 12.0 seconds and the probable error of the difference 1.8 seconds. Since the difference is 6.7 times its probable error, it is reliable statistically.

TABLE 1

THE MEAN TIME IN SECONDS FOR NAMING FIFTY SYMBOLS

Group	Number of Subjects	Mean	$P.E._{dist.}$	Difference between Means	$P.E._{diff.}$
Control	20	40.50	7.42	12.00	1.80
Experimental	20	28.50	3.10		

These results show that the reaction time to stimuli is much shorter when one response tendency has been established than when several kinds of responses have been practiced. Since the conditions of this experiment seem to conform to those postulated by Peterson in his theory concerning the difference in time between naming colors and reading color names, the results secured are interpreted as supporting Peterson's explanation.

EXPERIMENT II

Problem. How will the difference between the time for naming objects and that for reading names of objects compare with the difference between the time for naming colors and that for reading the names of colors?

Materials. Red, blue, green, brown and purple were the colors used on the color charts. They were printed from 24 point swastikas, ten to the line, five lines to the half sheet. They were arranged to avoid any regularity of recurrence. Each color appeared twice in each line. Two orders were used. The test sheets for speed in reading names of colors were prepared by replacing each color in the color chart by its typewritten name, with the names spaced to give the same length of line.

The 'object naming test' comprised small pictures of the

following five common objects which were cut from the Kuhlmann-Anderson intelligence test for the first semester of the first grade: a watch, a chair, a hat, a wagon, and a basket. The pictures were pasted on a card after the same arrangement as used for the colors on the color test just described. (Each of the five colors was replaced by one of the pictures, *i.e.,* wherever the color red was located in the color test, it was replaced by the picture of a hat, the color blue was replaced by the picture of a chair, etc.) The name reading test was made by typing the names of the objects in the same order as that used in the picture sheet.

Subjects and Procedure. Fifty sophomore students of general psychology at David Lipscomb College in the fall quarter of 1937 served as subjects. Two cards of 50 pictured objects were named and two sheets of 50 names of objects were read by each subject. The double fatigue order was followed. Twenty-five subjects, selected at random, followed the experimental order: named objects, read words, read words, named objects; and the other twenty-five subjects read words, named objects, named objects, and read words. The procedure employed in starting, checking time, etc. was the same as that followed in analogous tests in Experiment I. The color test and the names of colors test were administered in the same manner to the same subjects.

Results. The lower part of Table 2 shows that the difference between the time for naming objects and that for reading names of objects is 4 seconds greater than the difference between the time for naming colors and that for reading names of colors. This difference is statistically reliable, being 6.8 times its probable error. Since the subjects presumably were just as familiar with the objects as with the colors why should calling the names

TABLE 2

COMPARISONS OF SPEED IN NAMING AND READING RESPONSES FOR OBJECTS AND COLORS

Kind of Activity	Objects		Colors		The Difference between the Differences, and Its Probable Error
	Mean	$P.E._{dist.}$	Mean	$P.E._{dist.}$	
Naming	34.5	2.75	28.9	2.39	
Reading	21.8	1.46	20.2	1.56	
Difference	12.70		8.70		4.00
$P.E._{diff.}$	.44		.40		.60

of objects require more time than calling the names of colors? The writer suggests that the greater interference met in the serial naming of objects may be due to the fact that more specific tendencies to *non-verbal* action have been established toward these common objects or stimuli than toward colors, which are special, 'detachable' aspects of objects.

Experiment III

Problem. How does the mean time for sorting cards according to color compare with that for sorting cards according to names of colors?

Material. A deck of 50 cards was used at each sorting. For sorting according to word, each of the five color names was equally represented in the deck. For sorting by color each of the five colors was equally represented in the deck. The cards were arranged in the deck to avoid the repetition of the same word or color in succession. The cards used were slightly larger and slightly heavier than playing cards. The words were printed from 18 point Franklin lower case type and the colors from 24 point swastikas. The compartments into which the cards were sorted were arranged in a single row and were labelled as follows from left to right: red, blue, green, brown, and purple. The same order was maintained throughout the experiment. When the sorting was according to color names, color names were used as labels on the compartments; when according to color, colors were used as labels. All errors in sorting were corrected by recovering cards that were incorrectly placed and putting them in the right compartments.

Subjects and Procedure. The subjects were 40 college undergraduates (20 males and 20 females). As a preliminary exercise to facilitate handling the cards, each subject dealt a deck of plain cards in the ordinary way. Each subject then sorted six decks of cards. Half of the subjects (10 males and 10 females) sorted the six decks in the following order: colors, words, words, colors, colors, words; and the other (10 males and 10 females) reversed the order. In all cases the scores used were those of the last four sortings, allowing the first two sortings for practice in locating compartments and for adjustment to the task.

Results. The results given in Table 3 show that the mean time for sorting cards according to colors is less than that for sorting cards according to words (names of colors). As shown in the last column of the table, however, the difference is not statistically reliable. The difference of 3.7 seconds found in the female group is considered significant since the time for sorting colors was less than that for sorting words, in 15 of the 20 cases, and equal to it in 2 of the remaining 5 cases. Attention is also called to the fact that the difference between the sex groups in sorting words is only 1.3 seconds while in sorting colors it is 5.3 seconds, a difference which is 3.73 times its probable error. Part of the difference between the sex groups is accounted for by the fact that several of the men, although not color blind, stated that they had some difficulty in distinguishing some of the colors. But for this fact, the difference in speed between the two types of sorting would have been considerably greater than that reported in Table 3.

TABLE 3

A COMPARISON OF THE MEAN TIMES FOR SORTING FIFTY CARDS ACCORDING TO WORDS (COLOR NAMES) AND FIFTY CARDS ACCORDING TO COLOR

Sex	Number of Subjects	Words Mean	Words $P.E._{dist.}$	Colors Mean	Colors $P.E._{dist.}$	Diff.	$Diff./P.E._{diff.}$
Males	20	49.80	4.65	50.10	5.05	1.30	.20
Females	20	48.50	4.76	44.80	4.09	3.70	2.66
M & F	40	49.40	4.72	47.60	4.86	1.80	1.80

These figures show that the reaction time for sorting colors tends to be shorter than that for *sorting* names of colors, which is just the reverse of the result found in the case of the *verbal* reactions (naming colors and reading color names). This reversal of speed relations indicates that the difference between time for naming colors and that for reading color names cannot be accounted for by the difference in the nature of the stimuli, or in the visual or recognitive processes as such. Instead it seems plausible to assume, with Peterson, that the greater speed in reading the printed color names results from the fact that this type of response is highly specific and has therefore been much more intensively practiced than the heterogeneous types of reactions to colors.

Summary and Conclusions

The purpose of the present study was to test by experiment Peterson's explanation of the fact that naming colors requires more time than reading the printed names of colors. It was his view that the latter response is a specific and highly practiced one, whereas the reactions to colors are diverse in nature, varying with the type of relation which the subject might sustain with reference to colored objects (perceiving, grasping and eating, admiring, etc.). Three experiments were devised to test this theory:

(1) In the first experiment one of two equivalent groups of subjects was given practice in the naming response to five nonsense symbols, and the other group was given practice in naming, sorting, collecting, tapping, and checking responses to the same five symbols. The total number of responses made by each member of each group was 1200. The difference between the mean for the two groups was 12.0 seconds, and the probable error of the difference was 1.8. Since this difference is statistically reliable it indicates that the reaction time to stimuli is shorter when one response habit has been established than when several different response tendencies have been developed.

(2) In the second experiment 50 sophomores were given the following tests: (a) naming five common objects (100 reactions to random arrangements of the pictures); (b) reading the printed names of the same objects (100 reactions); (c) naming five colors (100 reactions); (d) reading the printed names of the same colors (100 reactions). The difference in average time between (a) and (b) was found to be 4 seconds greater than the difference between the times for (c) and (d), and since the probable error of this value is only .595, it is highly reliable statistically. It is assumed that the disparity between the results for objects and those for colors is due to the fact that people normally develop more specific non-verbal response tendencies to common objects than to colors, which represent special 'detachable' aspects of all sorts of objects.

(3) The third experiment compared the time for *sorting* a deck of cards according to color with the time for sorting cards

bearing the printed names of the same colors. Contrary to the result when the corresponding types of verbal reactions were compared, the average time for sorting the colored cards was less than that for sorting the cards with printed color names. Although this difference was not reliable statistically, the contrast between the naming and the sorting reactions is nevertheless significant. It would seem that the comparative slowness in sorting the cards with printed words is due to the interference effect consequent upon being forced to make *manual* responses to stimuli which, in the case of most people, usually evoke *reading* reactions. Since normal reactions to colors are heterogeneous in nature, and include naming reactions among many other types, the interference effects of a highly specific response pattern would not be operative.

The results of all three experiments would thus seem to support Peterson's explanation of the difference in time between color naming and word reading reactions. In addition to providing direct experimental evidence in support of this hypothesis, it would seem that the experimental procedures used are susceptible of application to the study of many conditions which affect motor performance, particularly to problems related to facilitation and inhibition.

Bibliography

1. Brown, Warner. Practice in associating color names with colors. *Psychol. Rev.*, 1915, **22**, 45–55.
2. Cattell, J. McK. The time it takes to see and name objects. *Mind*, 1886, **11**, 63–65.
3. Garrett, H. E., and Lemmon, V. W. An analysis of several well-known tests. *J. Appl. Psychol.*, 1924, **8**, 424–438.
4. Ligon, E. M. A genetic study of color naming and word reading. *Amer. J. Psychol.*, 1932, **44**, 103–121.
5. Lund, F. H. The rôle of practice in speed of association. *J. Exper. Psychol.*, 1927, **10**, 424–433.
6. Peterson, J., Lanier, L. H., and Walker, H. M. Comparisons of white and negro children. *J. Comp. Psychol.*, 1925, **5**, 271–283.
7. Stroop, J. Ridley. Interference in serial verbal reactions. *J. Exper. Psychol.*, 1935, **18**, 643–661.
8. Stroop, J. Ridley. The basis of Ligon's theory. *Amer. J. Psychol.*, 1935, **47**, 499–504.
9. Woodworth, R. S., and Wells, F. L. Association tests. *Psychol. Rev. Monog. Suppl.*, 1911, **13** (No. 57). Pp. 85.

MATURITY AND LEARNING ABILITY

Wendell W. Cruze
WILSON TEACHERS COLLEGE

Within the last few years, several articles dealing with the relationship between age and learning ability have appeared. Most of these have dealt with the learning ability of lower organisms. Watson (13) obtained results which indicated that young rats could learn to run a maze more readily than adult rats, though the adult rats were superior in learning problems which depended upon control of movement for their solution. Yerkes (14), on the other hand, found that the 10 months old dancing mice could learn to run a maze more readily than mice 1 or 2 months of age; and black and white discrimination habits were formed more readily in the younger mice. Hubbert (6) working with five groups of rats, with respective ages of 25, 65, 200, 300, and 500 days, concluded that the ability to learn to run mazes decreased with age. Liu (7) using seven groups of rats, with respective ages of 30, 45, 60, 75, 100, 150, and 250 days, concluded that the ability of the rat to learn to run a maze increased rather rapidly from 30 to 75 days of age and then slowly decreased up to 250 days of age. The optimal learning period corresponded rather closely to the average age of ovulation.

Stone (10, 11) has published results which indicate that some of the results mentioned above were obtained because the factor of motivation was neglected. In his experimental work he used five groups of rats with ages of 30 days, 50 days, 100 days, 8 months, and 2 years, respectively. The hunger drive was the type of motivation which was used throughout, and this factor was carefully controlled. He used both problem boxes and mazes as learning problems. His results indicated that the maximum learning rate was reached between 30 and 75 days, and that the learning ability of rats does not decline during the first two thirds

of their lives. Stone pointed out that different results might be obtained if unequal motivation at the different ages were employed, or if a problem especially suited to the behavior tendencies of a certain age were used.

It is interesting to note that in all of these experiments the first few days of the life of the animals have been neglected. Another line of investigation, however, has dealt with this part of the life of the chick. Several studies pertaining to the development of accuracy in the pecking reaction have appeared (1, 2, 3, 5, 8, 9). In general the results obtained have indicated that accuracy in the complete pecking reaction (striking, seizing, and swallowing) developed more rapidly when the initial practice in this reaction was delayed from 2 to 5 days after hatching than when practice was allowed immediately after hatching or within a period of 24 hours after hatching. Some investigators have interpreted these results to indicate that this increased accuracy was due to specific maturation of the nervous system, while others have pointed out the similarity existing between the curves of achievement in this function and other curves representing the formation of motor habits. In other words, some investigators would use the word "maturation" to explain this increased rate of development in accuracy, while others would say that "learning" is the true explanation. Anyone who is familiar with the early life of a chick will testify to the fact that the chick does mature to some extent, and in various ways, during the first few days of its life. This maturation is generally observed as increased ability to support the body, better coördination, and a general increase in bodily activity.

The Present Problem

In an attempt to bridge the gap between these two lines of research, the experiments reported in this paper have been carried out. An effort has been made to ascertain whether or not the maturation which takes place with increasing age in the early days of the life of the chick will bring about an increase in the ability of the chicks to learn to run mazes and to solve problem boxes. A comparison of the achievements of chicks of

various ages, when factors other than general maturation have been eliminated or experimentally controlled, is one way of attacking this problem and is the technique followed in this study.

The Subjects and Experimental Procedure

All of the chicks used in this study were Barred Plymouth Rocks and came from blood-tested flocks. In each problem six groups of chicks were used. One group was started on the problem when 24 hours old, another group at the age of 48 hours, another at 72 hours, two groups at 96 hours of age, and the oldest group started on the problem when 3 weeks of age.

It seemed possible that the strength of the tendency to go toward and be with other chicks might increase with age, not as a result of maturation but as a result of constant association with the other chicks. If this were true, and no attempt were made to measure or control this motive, the results of the experiments would be vitiated to some extent. It would be impossible to tell to what extent the results were due to maturation and to what extent they were due to social motivation. In an attempt to check on the possible effects of social motivation forty-nine chicks were raised in isolation from other chicks. They were hatched in the dark room of the laboratory and were taken, one at a time, to separate compartments in a lighted room. Electric brooder facilities were available in each compartment so that each chick could be kept at the proper temperature. In this way, it was possible to raise chicks to the age of ninety-six hours that had never seen other chicks. It was impossible, however, to keep them from hearing the chirping of other chicks. After these chicks solved the various problems the first time, they were allowed to remain with the other chicks. In all other respects, these chicks were cared for in the same way as all the other chicks.

The chicks which were raised in a normal environment were taken from the commercial hatchery incubator immediately upon hatching and were removed to the laboratory at once. This procedure provided for a period of at least 24 hours during

which the chicks were allowed to become adjusted to their surroundings. In the laboratory they were kept in a feeding pen at one end of which was a brooder. In this pen these chicks were kept until they reached the age at which they were to start their tests. At night all food and water were removed from the pen. The chicks were taken from the pen for their tests before being fed, and since food was before them for only 12 hours each day this meant that they had been without food for 12 hours. The food used throughout the experiments was a commercial starting food, Purina Startina, which contained cod liver oil.

An attempt was made to hold factors such as temperature, light conditions, and motivation constant. In order to control seasonal variations in temperature, an electric heater was used. When the room was cold, the heater was placed near the experimental table, and as the room temperature increased the distance between the heater and the experimental table was increased. In this way, it was possible to hold the temperature around 80 to 85° Fahrenheit during the experimental periods. As a further control of possible temperature and seasonal effects, approximately the same number of chicks from each hatching was included in each group. Since artificial lighting was used throughout, it was easy to hold this factor relatively constant.

Every effort was made to hold the factor of hunger motivation constant.[1] Each chick ran the maze of the problem box in the morning when it had been without food for about 12 hours. During the other 12 hours out of each 24-hour period, the chicks were allowed a limited amount of food and plenty of water. The food was placed in the pen at intervals of 3 hours during the 12 hour period. Eight grams of food were allowed for each chick each day. This amount, which is slightly below the normal food consumption for chicks of this age,[2] was decided

[1] The results obtained with the "Isolated" groups indicated that the factor of social motivation was operating in all of the other groups. A consideration of the possible effects of this factor is included in the discussion of results.

[2] A preliminary feeding experiment indicated that chicks of this breed ate an average of 9 grams of Purina Startina a day during the first three weeks of life, when food was kept before them all of the time.

upon in order to insure strong hunger motivation. The close correspondence of the weight records of the different groups, shown below in Table 1, indicates that this method was successful in keeping the motivation factor at approximately the same

TABLE 1

AVERAGE WEIGHT IN GRAMS OF EIGHTEEN GROUPS OF CHICKS

		Age				
		1 day	5 days	10 days	15 days	20 days
Group A	Mean	36.14	36.38	38.91	49.26	63.42
	S.D.	2.16	2.42	4.08	6.42	10.15
Group B	Mean	35.96	37.14	41.28	50.63	61.62
	S.D.	1.98	2.08	3.69	5.98	8.43
Group C	Mean	36.32	36.15	38.96	51.42	60.09
	S.D.	3.01	2.97	3.09	4.86	7.36
Group D	Mean	37.39	38.21	41.62	53.61	66.72
	S.D.	2.56	2.68	5.01	8.43	11.26
Group E	Mean	38.61	37.69	41.83	52.09	63.42
	S.D.	2.75	2.61	4.32	7.21	10.84
Group F	Mean	37.63	38.06	42.38	53.69	67.47
	S.D.	1.86	1.91	3.69	6.26	10.15
Group G	Mean	36.94	38.02	42.61	51.96	64.15
	S.D.	2.96	2.49	5.01	7.36	10.29
Group H	Mean	38.21	37.68	43.19	52.45	69.12
	S.D.	3.25	3.28	5.15	8.02	12.23
Group I	Mean	37.33	37.96	41.06	49.43	62.61
	S.D.	2.78	3.02	4.19	6.79	10.09
Group J	Mean	37.86	36.91	39.36	47.88	58.62
	S.D.	3.06	3.15	4.71	6.63	9.03
Group K	Mean	35.98	36.72	39.35	49.52	61.29
	S.D.	2.09	2.18	3.99	6.09	9.56
Group L	Mean	36.46	37.34	41.68	53.71	65.02
	S.D.	1.96	2.07	4.32	7.28	11.81
Group M	Mean	36.86	37.02	40.15	52.43	64.15
	S.D.	2.69	2.48	4.39	6.43	8.67
Group N	Mean	37.73	37.46	39.98	51.16	61.39
	S.D.	3.18	3.21	5.06	7.08	9.39
Group O	Mean	36.50	37.00	40.76	50.98	66.29
	S.D.	1.98	1.85	4.15	6.56	11.02
Group P	Mean	37.32	35.98	39.03	52.15	65.43
	S.D.	2.43	2.26	4.65	7.14	10.29
Group Q	Mean	38.18	38.19	40.96	49.14	61.48
	S.D.	3.14	3.12	5.14	7.26	9.14
Group R	Mean	36.81	37.02	41.24	53.52	68.92
	S.D.	2.08	2.16	5.08	8.27	11.86

strength in the different groups. When a chick died, its weight record was discarded from the very beginning.

In these experiments, three problems were used: Two mazes and a problem box. A brief description of each is given and the dimensions are indicated on the accompanying diagrams.

Plan for "S" Maze

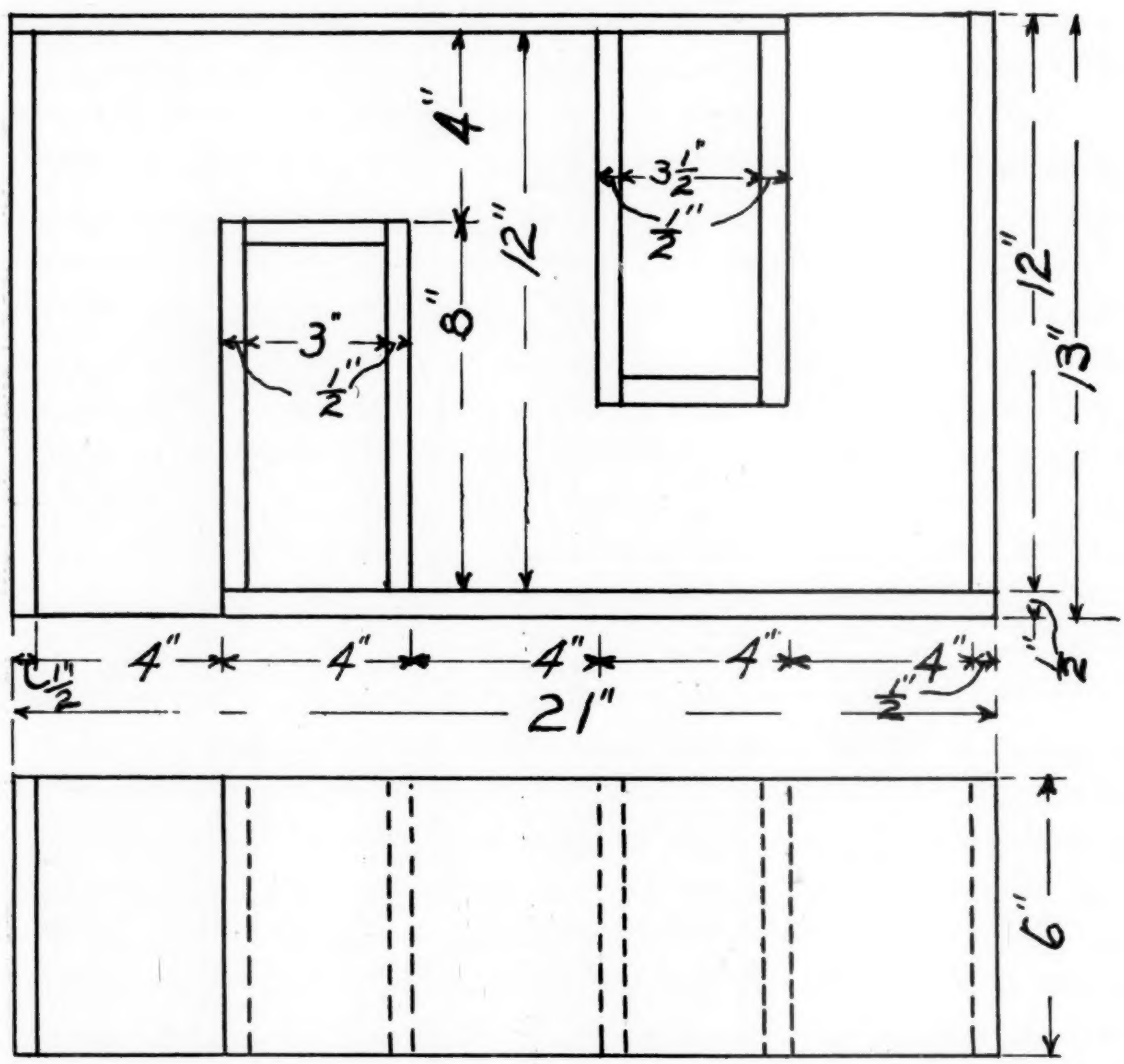


Figure 1. The plan of the "S" maze, showing dimensions of the floor plan (above) and of the walls (below).

1. *The "S" Maze.* The walls of the "S" maze (Fig. 1.) were constructed of wood and the floor of metal. The maze was covered with ½ inch hardware cloth and was painted a flat black inside and out. Each chick was carried to the entrance of the maze in a small box of the same material as the maze. The

gate of the box was then opened, allowing the chick to enter the maze. The time consumed in traveling through the maze to the decoy chicks [3] was the score. The stop watch was started when the gate of the box was opened and stopped when the chick jumped down into the pen with the decoy chicks. The pen was 3 inches below the level of the maze floor. This was not high enough to cause the chicks to hesitate in jumping down, but it served the purpose of keeping the decoy chicks out of the maze. The reliability of the maze was calculated by correlating the time scores made on the odd trials with the time scores made on the even trials. The Spearman-Brown formula was used to change the coefficients of correlation into coefficients of reliability, and the results are given in Table 2.

TABLE 2

COEFFICIENTS OF RELIABILITY FOR "S" MAZE

Group	Coefficient of Reliability
24-hour group	.80
48-hour group	.69
72-hour group	.72
96-hour group	.82
3-week group	.75
Isolated group	.87

2. *The Multiple "T" Maze.* This was a very simple multiple "T" maze (Fig. 2) consisting of two sections with two blind alleys. The true path was the right-hand path at each bifurcation. The maze was constructed of wood, with a tin floor, and was covered with ½ inch hardware cloth. It was painted a flat black inside and out. Each chick was carried to the entrance of the maze in a small box of the same material as the maze. The gate of the box was then opened, allowing the chick to step into the maze. The stop watch was started when the gate of the box was opened and stopped when the chick jumped down into the pen with the decoy chicks. As in the case with the "S" maze, the pen was three inches below the level of the maze floor. The reliability of the maze was calculated in the manner just described

[3] Ten chicks, of the same age as the experimental chicks, were used as a decoy group. The decoy chicks were actively chirping and eating during the experiments.

for the "S" maze. The reliability coefficients are given in Table 3.

PLAN FOR "T" MAZE

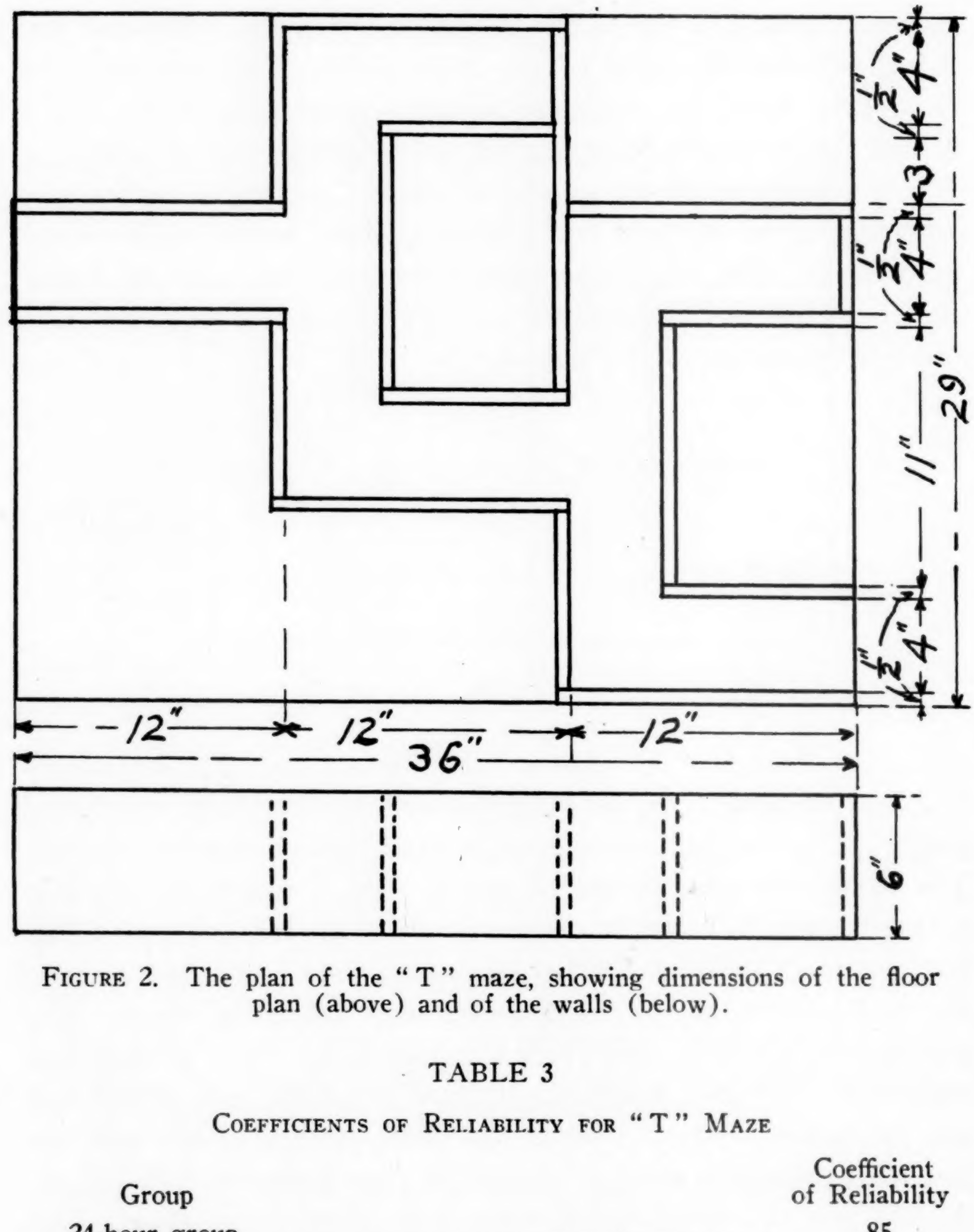


FIGURE 2. The plan of the "T" maze, showing dimensions of the floor plan (above) and of the walls (below).

TABLE 3

COEFFICIENTS OF RELIABILITY FOR "T" MAZE

Group	Coefficient of Reliability
24-hour group	.85
48-hour group	.76
72-hour group	.86
96-hour group	.89
3-week group	.84
Isolated group	.90

3. *The Problem Box.* The problem box (Fig. 3), a modification of the one used by Thorndike (12), was constructed of wood with a floor of metal and was painted a flat black inside and out. The front side of the box was made of ordinary screen wire through which the chicks could see but not pass. The only way out of the box was by means of an inclined plane, made of wood

PROBLEM BOX

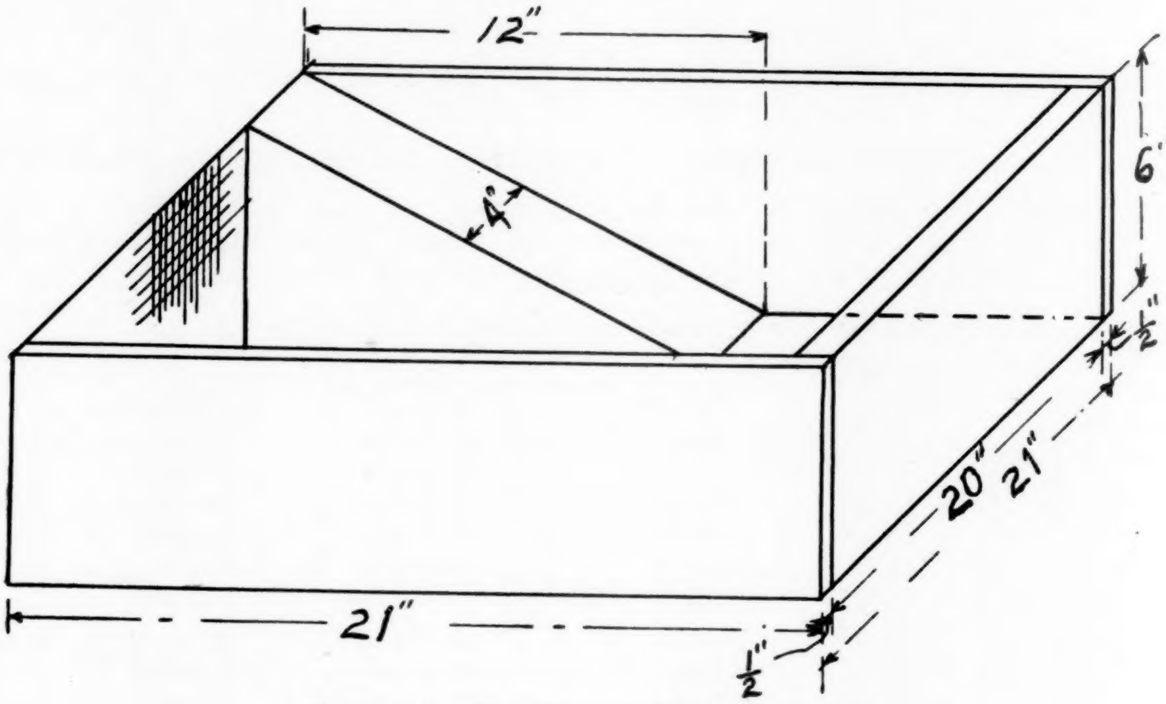


FIGURE 3. The plan of the problem box.

and covered with ½ inch hardware cloth to prevent the chicks from slipping, which led from the floor of the box to the top of the screen on the front side. Each chick was placed in the box in the center, about three inches from the screen wire, and facing the decoy chicks which were feeding on the other side of the screen wire. A covering of ½ inch hardware cloth was immediately placed over the box, leaving an opening over the inclined plane. The stop watch was started when the chick was placed on the floor of the box and stopped when the chick jumped from the top of the inclined plane into the feeding pen with the decoy chicks. This pen was on the same level as the floor of the box. The reliability of the problem box was calculated by correlating the time scores made on the odd trials with the time scores made on the even trials. The Spearman–Brown formula was used to

change the coefficients of correlation into coefficients of reliability, and the results are given in Table 4.

TABLE 4

Coefficients of Reliability for the Problem Box

Group	Coefficient of Reliability
24-hour group	.70
48-hour group	.73
72-hour group	.73
96-hour group	.76
3-week group	.75
Isolated group	.81

Table 5 gives a summary of the facts concerning the manner in which each group of chicks was used.

TABLE 5

Experimental Conditions for Eighteen Groups of Chicks

Group	No. in Group	Problem	Age at Start	Environment
A	25	S Maze	24 hours	Normal
B	25	S Maze	48 hours	Normal
C	25	S Maze	72 hours	Normal
D	25	S Maze	96 hours	Normal
E	25	S Maze	3 weeks	Normal
F	18	S Maze	96 hours	Isolated
G	25	T Maze	24 hours	Normal
H	25	T Maze	48 hours	Normal
I	25	T Maze	72 hours	Normal
J	25	T Maze	96 hours	Normal
K	23	T Maze	3 weeks	Normal
L	15	T Maze	96 hours	Isolated
M	25	Problem Box	24 hours	Normal
N	25	Problem Box	48 hours	Normal
O	25	Problem Box	72 hours	Normal
P	25	Problem Box	96 hours	Normal
Q	25	Problem Box	3 weeks	Normal
R	16	Problem Box	96 hours	Isolated

Results

The amount of time and the number of trials necessary to learn the problems are given in Tables 6 and 7 in terms of means and standard deviations. The criterion of learning, arbitrarily selected, was the same in all of the problems, namely, 4 solutions out of 5 trials, in less than 30 seconds each. In order to present detailed comparisons of each age group with every other, for all test criteria, Tables 8–13 have been prepared.

TABLE 6

SHOWING COMPARISONS OF CHICKS OF DIFFERENT AGES, IN TERMS OF THE NUMBER OF TRIALS NECESSARY TO SOLVE THE THREE PROBLEMS

Problem		Means and Standard Deviations for Chicks of Different Ages					
		24-hour	48-hour	72-hour	96-hour	3-week	Isolated (96-hours)
"S" Maze	Mean	7.64	7.00	6.64	5.80	5.24	6.56
	S.D.	1.02	0.94	0.62	0.57	0.71	0.76
"T" Maze	Mean	9.32	8.48	7.72	7.00	6.60	7.80
	S.D.	0.93	0.90	0.92	1.10	1.05	0.89
Problem Box	Mean	8.32	6.92	6.48	5.32	5.28	6.37
	S.D.	1.11	1.01	0.86	1.26	1.08	0.93

TABLE 7

SHOWING COMPARISONS OF CHICKS OF DIFFERENT AGES, IN TERMS OF THE NUMBER OF SECONDS NECESSARY TO SOLVE THE THREE PROBLEMS

Problem		Means and Standard Deviations for Chicks of Different Ages					
		24-hour	48-hour	72-hour	96-hour	3-week	Isolated (96-hours)
"S" Maze	Mean	1083	577	405	312	138	696
	S.D.	511	331	269	187	96	204
"T" Maze	Mean	1508	1289	732	463	159	946
	S.D.	503	423	378	201	102	356
Problem Box	Mean	1251	629	579	340	148	680
	S.D.	634	372	384	210	113	323

TABLE 8

SHOWING THE STATISTICAL RELIABILITY OF THE DIFFERENCES OBTAINED WITH CHICKS OF DIFFERENT AGES IN TERMS OF THE NUMBER OF TRIALS NECESSARY TO SOLVE THE "S" MAZE. THE ADVANTAGE LIES ALWAYS WITH THE SECOND GROUP LISTED

Groups	D/σ_D
24-hour group and 48-hour group	2.29
24-hour group and 72-hour group	4.17
24-hour group and 96-hour group	8.00
24-hour group and 96-hour, isolated group	4.00
24-hour group and 3-week group	9.60
48-hour group and 72-hour group	1.57
48-hour group and 96-hour group	5.22
48-hour group and 96-hour, isolated group	1.63
48-hour-group and 3-week group	7.33
72-hour group and 96-hour group	4.94
72-hour group and 96-hour, isolated group	0.36
72-hour group and 3-week group	7.37
96-hour, isolated group and 96-hour group	3.63
96-hour group and 3-week group	3.11
96-hour, isolated group and 3-week group	5.71

These tables contain 'critical ratios' (Diff./S.D. of Diff.), and a difference is considered to be statistically reliable when this ratio is as great as 3.00. It is at once apparent that the older chicks are much better learners, as far as these specific problems are concerned, than the younger chicks. Omitting, for the time

TABLE 9

SHOWING THE STATISTICAL RELIABILITY OF THE DIFFERENCES OBTAINED WITH CHICKS OF DIFFERENT AGES IN TERMS OF THE AMOUNT OF TIME NECESSARY TO SOLVE THE "S" MAZE. THE ADVANTAGE LIES ALWAYS WITH THE SECOND GROUP LISTED

Groups	D/σ_D
24-hour group and 48-hour group	4.18
24-hour group and 72-hour group	5.89
24-hour group and 96-hour group	7.07
24-hour group and 96-hour, isolated group	3.41
24-hour group and 3-week group	9.09
48-hour group and 72-hour group	2.02
48-hour group and 96-hour group	3.49
96-hour, isolated group and 48-hour group	1.45
48-hour group and 3-week group	6.36
72-hour group and 96-hour group	1.43
96-hour, isolated group and 72-hour group	4.04
72-hour group and 3-week group	4.68
96-hour, isolated group and 96-hour group	6.30
96-hour group and 3-week group	4.14
96-hour, isolated group and 3-week group	10.73

TABLE 10

SHOWING THE STATISTICAL RELIABILITY OF THE DIFFERENCES OBTAINED WITH CHICKS OF DIFFERENT AGES IN TERMS OF THE NUMBER OF TRIALS NECESSARY TO SOLVE THE "T" MAZE. THE ADVANTAGE LIES ALWAYS WITH THE SECOND GROUP LISTED

Groups	D/σ_D
24-hour group and 48-hour group	3.44
24-hour group and 72-hour group	6.23
24-hour group and 96-hour group	8.03
24-hour group and 96-hour, isolated group	5.00
24-hour group and 3-week group	9.82
48-hour group and 72-hour group	3.16
48-hour group and 96-hour group	5.44
48-hour group and 96-hour, isolated group	2.29
48-hour group and 3-week group	7.00
72-hour group and 96-hour group	2.54
96-hour, isolated group and 72-hour group	0.41
72-hour group and 3-week group	4.19
96-hour, isolated group and 96-hour group	2.59
96 hour group and 3-week group	1.40
96-hour, isolated group and 3-week group	3.91

being, the records made by groups F, L, and R, which were raised in isolation, and which were therefore presumably working under different conditions of motivation, we get the following results. In about 45% of the cases an increase in age of 24 hours produced a reliable difference in both time and error

TABLE 11

SHOWING THE STATISTICAL RELIABILITY OF THE DIFFERENCES OBTAINED WITH CHICKS OF DIFFERENT AGES IN TERMS OF THE AMOUNT OF TIME NECESSARY TO SOLVE THE "T" MAZE. THE ADVANTAGE LIES ALWAYS WITH THE SECOND GROUP LISTED

Groups	D/σ_D
24-hour group and 48-hour group	1.66
24-hour group and 72-hour group	6.16
24-hour group and 96-hour group	8.93
24-hour group and 96-hour, isolated group	4.13
24-hour group and 3-week group	13.10
48-hour group and 72-hour group	4.89
48-hour group and 96-hour group	7.94
48-hour group and 96-hour, isolated group	2.74
48-hour group and 3-week group	14.65
72-hour group and 96-hour group	2.77
96-hour, isolated group and 72-hour group	1.80
72-hour group and 3-week group	7.38
96-hour, isolated group and 96-hour group	4.43
96-hour group and 3-week group	4.75
96-hour, isolated group and 3-week group	8.37

TABLE 12

SHOWING THE STATISTICAL RELIABILITY OF THE DIFFERENCES OBTAINED WITH CHICKS OF DIFFERENT AGES IN TERMS OF THE NUMBER OF TRIALS NECESSARY TO SOLVE THE PROBLEM BOX. THE ADVANTAGE LIES ALWAYS WITH THE SECOND GROUP LISTED

Groups	D/σ_D
24-hour group and 48-hour group	4.86
24-hour group and 72-hour group	6.54
24-hour group and 96-hour group	8.82
24-hour group and 96-hour, isolated group	5.76
24-hour group and 3-week group	10.52
48-hour group and 72-hour group	1.62
48-hour group and 96-hour group	4.97
48-hour group and 96-hour, isolated group	1.72
48-hour group and 3-week group	6.07
72-hour group and 96-hour group	3.84
72-hour group and 96-hour, isolated group	0.42
72-hour group and 3-week group	4.88
96-hour, isolated group and 96-hour group	2.89
96-hour group and 3-week group	0.19
96-hour, isolated group and 3-week group	3.47

TABLE 13

SHOWING THE STATISTICAL RELIABILITY OF THE DIFFERENCES OBTAINED WITH CHICKS OF DIFFERENT AGES IN TERMS OF THE AMOUNT OF TIME NECESSARY TO SOLVE THE PROBLEM BOX. THE ADVANTAGE LIES ALWAYS WITH THE SECOND GROUP LISTED

Groups	D/σ_D
24-hour group and 48-hour group	4.23
24-hour group and 72-hour group	4.51
24-hour group and 96-hour group	6.80
24-hour group and 96-hour, isolated group	3.78
24-hour group and 3-week group	8.54
48-hour group and 72-hour group	0.47
48-hour group and 96-hour group	3.40
96-hour, isolated group and 48-hour group	0.46
48-hour group and 3-week group	6.17
72-hour group and 96-hour group	2.72
96-hour, isolated group and 72-hour group	0.90
72-hour group and 3-week group	5.39
96-hour, isolated group and 96-hour group	3.74
96-hour group and 3-week group	4.00
96-hour, isolated group and 3-week group	6.33

scores. In all of the cases an increase of 48 hours in age produces a reliable difference in the time scores and in all except 2 cases (the 96-hour group and the 3-week group on the "T" maze and the same groups on the problem box) a reliable difference in the trial scores. It is possible that the physiological limit of learning, in terms of the number of trials, had been approached in these 2 cases. Further increases in age differences served to increase the reliability of the differences in scores.

There are 2 factors which seem to explain this increase in the rate of learning. They are (1) the development of a tendency to be with other chicks which is referred to as the social motive, and (2) general physiological maturity.

The groups of chicks raised in isolation were introduced in an effort to study the influence of social motivation. It is apparent from the results that the social motive does contribute to the drive toward the solution of the problem. The chicks in groups D, J, and P, (96 hours of age) that had formed the habit of being with other chicks, solved each problem in less time than the chicks in groups F, L, and R, chicks of the same age but in which the social motive had not been developed. These differences in time scores on all 3 problems are statistically reliable.

When trial scores are used, it is found that the groups of chicks in which the social motive has been developed solve the problems in fewer trials. A statistically reliable difference is obtained with the " S " maze, and the differences obtained with the " T " maze and the problem box approach statistical reliability (D/σ_D equals 2.59 and 2.89 respectively).

It is quite probable that greater differences would have appeared if it had been possible to continue the isolation policy for a period longer than 96 hours. However, lack of facilities for keeping a large number of chicks in isolation for a long period of time prevented this.

It must not be supposed, however, that the development of this social motive was solely responsible for the increase in rate of learning. Reference to Tables 8–13 will show reliable differences in favor of the 96-hour isolated groups in all cases when the 24-hour groups are compared with the 96-hour isolated groups. These differences, which could not have been due to the development of the social motive since no opportunity was allowed for its development in the 96-hour isolated groups, were very probably due to the increase in general physiological maturity of the chicks. Comparisons of the 96-hour isolated group with the 48-hour group reveal that the former surpass the latter in all three problems by the 'trials' criterion, but on two of the three problems (S-maze and problem box) the 48-hour group excel the 96-hour isolated group in time required to learn the problems. Comparing the 96-hour isolated group with the 72-hour group, it is seen that the 'isolated' chicks surpass the 72-hour group in terms of trials, on two of the three problems (S-maze and problem box), but are inferior to them on all three problems when time is the basis of comparison. The majority of these differences are statistically unreliable, but nevertheless the general trend of the results is interesting. It would seem that to a considerable extent the 'social motive' consists in an increased sensitivity to the chirping of the decoy chicks, with the result that the rate of activity is considerably increased. But the comparison of the two 96-hour groups indicates that the effects of 'socialization' express themselves in the form of a reduction in

trials as well as in time, since the 'isolated' group is inferior to the normal group by both criteria, in all three tests. The 'heightened sensitivity' of the 'social' chicks is productive not merely of faster activity but of more rapid mastery of the problems, when age is controlled.

Summary and Conclusions

The attempt has been made to determine whether or not the maturation which takes place with increasing age in the early days of the life of the chick will bring about an increase in the ability of chicks to learn to run mazes and to solve problem boxes. A further aspect of the study has been the effort to ascertain whether or not social factors contribute to the increase in learning ability of normal chicks during this period. Three simple problems or tests were used: an S-maze, a T-maze and a problem box. The number of trials, and the time of each trial, constituted the learning records. The criterion of learning was four solutions out of five trials, in less than 30 seconds each.

Five groups of normal chicks were used, of the following ages: 24 hours, 48 hours, 72 hours, 96 hours, 3 weeks. By a 'normal' chick is meant one which was raised in a group of chicks of the same age and which displayed no symptoms of ill health. An additional group of chicks was studied, each member of which was isolated from other chicks from the time of hatching until 96 days of age, the time of the tests for this group.

The test results for the five normal groups indicated a progressive increase in learning ability with increase in age. In about 45% of the cases an increase of 24 hours in age produced a statistically reliable difference in favor of the older group, in terms of both time and errors. The differences were statistically reliable in all cases where the difference in age was 48 hours or more, except in the case of trials on T-maze and problem box for the comparisons of the 96-hour and 3-week groups. It is suggested that the simple problems were not adequate to produce a differentiation between these two age groups.

The results with the control or isolated group, indicate that mere maturation was not responsible for all of the increase in

proficiency with age. The 96-hour isolated group was inferior to the 96-hour normal group by all criteria, and four of the six differences were statistically reliable. When compared with the 72-hour group, the ' isolated ' chicks were inferior to the younger normal group on the basis of four of the six criteria and one of these differences (time on the S-maze) was reliable. The ' isolated ' chicks are relatively better as regards trials than in terms of time. This fact is further emphasized by the comparisons with the 48-hour group. The ' isolated ' chicks surpass the 48-hour chicks on all three tests in terms of trials, but are inferior to them in two of the three tests on the basis of time.

Bibliography

1. BIRD, C. The relative importance of maturation and habit in the development of an instinct. *Ped. Sem.*, 1925, **32**, 68–91.
2. ——— The effect of maturation upon the pecking instinct of chicks. *Ped. Sem.*, 1926, **33**, 212–234.
3. ——— Maturation and practice: Their effects upon the feeding reactions of chicks. *J. Comp. Psychol.*, 1933, **16**, 343–366.
4. CARMICHAEL, L. A re-evaluation of the concepts of maturation and learning as applied to the early development of behavior. *Psychol. Rev.*, 1936, **43**, 450–470.
5. CRUZE, W. W. Maturation and learning in chicks. *J. Comp. Psychol.*, 1935, **19**, 371–409.
6. HUBBERT, H. B. The effect of age on habit formation in the albino rat. *Behav. Monog.*, 1915, **2**. Pp. 55.
7. LIU, S. Y. The relation of age to the learning ability of the white rat. *J. Comp. Psychol.*, 1928, **7**, 75–86.
8. MOSELEY, D. The accuracy of the pecking response in chicks. *J. Comp. Psychol.*, 1925, **5**, 75–97.
9. SHEPARD, J. F., and BREED, F. S. Maturation and use in the development of an instinct. *J. Anim. Behav.*, 1913, **3**, 274–285.
10. STONE, C. P. The age factor in animal learning: I. Rats in the problem box and the maze. *Genet. Psychol. Monog.*, 1929 A, **5**, 1-130.
11. ——— The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze. *Genet. Psychol. Monog.*, 1929 B, **6**, 125–202.
12. THORNDIKE, E. L. *Animal Intelligence: Experimental Studies.* New York: Macmillan, 1911. Pp. viii+297.
13. WATSON, J. B. *Behavior: An Introduction To Comparative Psychology.* New York: Holt, 1914. Pp. xii+439.
14. YERKES, R. M. Modifiability of behavior in its relation to age and sex of the dancing mouse. *J. Comp. Neurol. & Psychol.*, 1909, **19**, 237–271.

THE BEHAVIOR OF RIGHT AND WRONG RESPONSES DURING WORK AND REST INTERVALS

KATHERINE VICKERY
ALABAMA COLLEGE

Since William James first stated that we "learn to skate in summer and swim in winter," numerous experiments have confirmed the value of an interval of rest for most learning tasks. Whenever practice has continued over a period of time, improvement in function has been evident. Under certain well defined conditions of spaced learning, improvement has been more rapid than when practice periods were continuous. Such improvement has been measured in terms of decrease in time and errors, or of increase in the amount learned within a given time. The length of the interval between practices, the nature of the learning material, the degree of progress in learning, the order of repetitions during the practice period, the manner of study, and the aim of learning are all important factors which have been experimentally investigated. But the relative behavior of right and wrong responses during continuous periods of practice and during intervals of rest has not been studied. The purpose of this experiment is to determine whether or not right responses tend to become fixed more rapidly during rest periods or during practice periods and whether wrong responses tend to be eliminated more rapidly during rest periods or during practice periods, or whether the reverse is true. If such phenomena exist, will they hold for all individuals and are they constant for varying periods of rest?

Most studies of massed and distributed learning tend to consider the interval of rest as a great advantage in the learning process. Summaries of studies of massed and distributed learning may be found in articles by Ruch (4) and Warden (5). A recent experiment by Lorge (1) in which numerous other factors

were controlled substantiates the findings of earlier investigators. On the other hand, experiments on the retention of material barely learned contradict the thesis that correct responses are strengthened during the period of disuse (or rest, as it is called in this experiment). They also furnish evidence that forgetting occurs at a maximum rate immediately after the learning period. These seemingly contradictory points of view may be due to the difference in the operations which define 'learning' and 'forgetting' in the two types of experiment. The two functions, learning and memory, are rather closely bound together. The present study does not attempt a reconciliation of points of view or an analysis of classifications. It purposes merely to study the behavior of correct and incorrect responses during the practice period and during varying periods of rest.

The Method of Procedure

A modification of the Peterson Rational Learning Test (2, 3) has been used in this study because it shows the relative reduction in errors and the increase in right responses during practice and during the period of rest. The test is described as one that measures "rational organization" and learning to "react to a changing situation." This type of learning is essentially different from the associative type involved in learning to use a typewriter. It deals with situations in which the factors do not remain constant throughout the problem. In the original test the learning problem consisted of associating the first ten letters of the alphabet with numbers from one to ten inclusive. The experimenter called the letters in order and the subject "guessed" the correct number. After the list of ten letters had been correctly "guessed" the order was changed and the series repeated until the subject could associate all the letters and their corresponding numbers through two successive series. The problem was then considered learned.

In the present experiment, twenty instead of ten letters were used and the subjects were required to associate these with numbers from one to twenty inclusive. The number of letters was increased to make the problem more difficult and to permit the

measurement of a number of different rest and practice periods. Eight random arrangements of the twenty letters were prepared and these eight arrangements were used in the same sequence for each subject. If more than eight trials were necessary for learning the problem, the subjects began again with the first random arrangement and followed the series of arrangements through as before. The following directions were read to the subjects before the experiment was begun:

"This is an experiment in rational learning. I am going to call out to you in succession twenty letters from 'a' to 't' inclusive in a random order. You are to continue guessing numbers for each letter called until I say 'Right,' when another letter will be called, and so on through the series. The numbers you guess are not to be above, but must include, twenty. Every error or wrong guess you make will be recorded in the test. We will go through the list twice at each sitting and these sittings will continue daily (every other day in the case of Group B and every Tuesday and Thursday in the case of Group C) until all letters and their corresponding numbers become associated so that you can go through the list twice in succession without error. Try not to keep the experiment in mind between trials. Do not ask questions after the experiment begins. Do you understand?"

In order to eliminate any feeling of competition or any necessity for extra effort on the part of subjects, which might cause them to work on the problem while not engaged in the practice period, they were told that their names would not be attached to their records, since the records were of value only as data for the class in experimental psychology. The subjects were divided into three separate groups, which were tested with varying intervals of rest but at approximately the same hour on each testing day. Group A was given two successive tests every day until all twenty letters and numbers became associated. This group was composed of seventeen subjects, of whom nine were college freshmen, two seniors, three graduate students, and three college professors. Eleven of the group were women and six were men. Group B was tested with two successive trials every other day. It was composed of eleven college women, members of a junior class in psychology. Group C worked on Tuesdays and Thursdays, giving one rest interval of two days and another of four days, alternately. It was composed of thirteen college women, of whom six were sophomores and seven juniors.

The Results

All errors on the tests were scored according to the plan used by Peterson (2). Errors were classified as: (1) unclassified, (2) logical, and (3) perseverative. Unclassified errors include, of course, all wrong responses. Logical errors are those made by guessing a number that had already been used for a letter occurring earlier in the series and which could not possibly be correct a second time. Perseverative errors are made by repeating an incorrect guess while reacting to the same letter. The sum of all three types of errors gives the error score. In the present study responses are scored 'right' whenever the subject gives correctly the number which corresponds with the letter without intervening incorrect guesses. The present situation only is considered. Should the subject later attempt to guess this number for other letters within the trial or in subsequent trials, the original 'correct' score is not affected. The number of correct responses in any one trial is then the number of times in each trial series that the subject calls the correct 'code' number on presentation of the letter without first guessing an incorrect number. Some element of guess-work is involved here, but after the first series is completed it is impossible to say to what extent the responses are due to 'chance' and to what extent to learning. Since a correct response in the first series is purely a matter of chance, except that logical and perseverative errors could be avoided, this series has been omitted in the final analysis of the results. The tables of scores, Tables 1, 2, 3, 4, 5, and 6, include the results of the first series, but this series was omitted when results of correct responses were computed. Table 1 gives the results in terms of errors for Group A, Table 2 gives the correct responses for the same group; Table 3 contains the error scores for Group B, and Table 4 the correct responses for this group; Table 5 gives the error scores for Group C, and Table 6 the correct responses for the same group.

To determine what happens to the errors during the period of practice, successive trials in the same practice period (*i.e.,* Trials

TABLE 1

ERROR SCORES OF GROUP **A**, COMPOSED OF 17 SUBJECTS WHO WERE GIVEN TWO TRIALS DAILY UNTIL LEARNING WAS COMPLETED

Trials / Subjects	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1.	173	170	177	221	171	86	71	64	37	16	27	9	16	6	4	4	5	1	0	2	2	3	1	0	0
2.	182	239	98	65	50	56	12	4	18	8	11	5	5	1	2	0	0								
3.	150	133	34	7	2	0	0																		
4.	267	300	182	52	30	19	3	0	1	0	2	1	0	0											
5.	179	176	136	120	141	115	107	73	56	67	35	17	13	6	6	3	2	2	0	1	1	0	0		
6.	179	172	169	151	129	123	106	94	87	61	50	33	21	12	6	4	4	1	0	0					
7.	171	162	170	129	136	99	93	87	82	43	41	27	20	10	4	1	2	0	0						
8.	239	219	132	99	95	59	62	64	46	35	25	11	12	6	6	2	3	1	0	1	0	0			
9.	218	215	326	176	266	178	119	48	11	32	2	7	0	0											
10.	256	223	196	77	171	193	49	16	63	23	37	4	3	2	0	0									
11.	179	146	39	31	6	6	1	0	0																
12.	187	250	294	144	144	54	48	14	8	1	3	0	0												
13.	155	152	127	128	124	46	91	35	30	21	11	20	8	31	7	1	0	0							
14.	173	201	203	129	131	48	72	30	22	15	9	11	3	6	3	1	0	0							
15.	169	200	217	143	131	52	66	18	16	7	5	9	2	24	6	3	1	0	0						
16.	190	248	289	139	136	50	46	19	5	0	2	1	0	0											
17.	184	250	299	148	146	58	50	9	11	1	4	0	0												

TABLE 2

CORRECT RESPONSES OF GROUP **A**, COMPOSED OF 17 SUBJECTS WHO WERE GIVEN TWO TRIALS DAILY UNTIL LEARNING WAS COMPLETED

Trials / Subjects	1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1.	3	2	3	4	9	7	9	10	17	14	13	17	16	17	16	17	17	19	20	19	18	18	20	20
2.	2	4	8	7	13	13	16	17	14	16	16	18	18	19	20	19	20	20						
3.	1	2	12	17	19	20	20																	
4.	2	1	6	10	16	18	19	20	19	29	19	19	20	20										
5.	1	2	4	8	7	9	13	14	15	14	15	16	16	18	19	18	19	19	20	19	19	20	20	
6.	1	1	3	4	8	7	8	10	11	12	14	16	17	17	19	18	19	19	20	20				
7.	3	2	3	5	6	9	8	10	12	11	14	15	15	16	18	19	19	20	20					
8.	2	3	8	9	11	11	14	14	15	13	15	16	15	16	18	18	19	19	20	19	20	20		
9.	3	1	2	2	1	6	9	11	16	16	15	19	16	20	20									
10.	2	0	2	5	3	7	8	15	8	12	9	18	17	19	20	20								
11.	1	6	12	14	16	18	19	20	20															
12.	1	3	5	9	8	10	11	16	18	19	18	20	20											
13.	3	4	4	3	4	7	6	12	12	13	15	14	18	14	18	19	20	20						
14.	2	1	1	4	5	10	8	11	12	14	16	18	19	18	19	19	20	20						
15.	2	0	0	5	6	10	9	16	17	18	18	17	19	17	18	19	19	20	20					
16.	1	2	4	6	8	8	11	15	16	20	20													
17.	1	4	3	8	9	12	14	16	16	17	17	18	18	18	19	19	19	20	20					

* Omitted in computation of results.

TABLE 3

Errors of Group B, Composed of 11 Subjects Who Were Given Two Trials Every Other Day Until Learning Was Completed

Trials / Subjects	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1.	200	183	94	92	16	5	3	0	2	1	0	2	0	0														
2.	336	260	229	253	225	195	213	163	134	144	73	91	33	52	13	0	8	3	0	0								
3.	489	293	153	47	27	25	43	4	10	6	0	0																
4.	152	165	204	193	111	118	76	17	4	0	0																	
5.	265	271	229	240	224	250	99	81	64	39	53	40	25	9	0	0												
6.	383	379	340	236	176	178	118	141	156	94	110	131	106	101	36	73	3	3	0	4	1	0	0					
7.	230	344	240	227	180	251	170	136	198	206	88	93	120	84	34	27	13	10	0	0								
8.	346	271	238	269	244	171	125	67	79	70	62	59	40	33	22	42	31	2	11	0	2	0	1	0	1	1	0	0
9.	123	285	311	219	180	145	140	84	89	149	84	137	2	1	2	0	0											
10.	463	389	164	213	120	215	91	76	68	67	45	37	0	2	1	2	0	0										
11.	232	233	155	186	216	159	245	286	173	214	237	213	40	37	0	0												

TABLE 4

CORRECT RESPONSES OF GROUP B, COMPOSED OF 11 SUBJECTS WHO WERE GIVEN TWO TRIALS EVERY OTHER DAY UNTIL LEARNING WAS COMPLETED

Trials / Subjects	1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1.	1	2	7	7	14	17	18	19	20	19	20	19	20	20														
2.	1	0	2	2	2	4	6	7	7	6	11	8	16	15	19	18	20	20										
3.	0	1	1	9	11	10	13	16	14	16	20	20																
4.	4	2	4	2	8	10	15	15	19	20	20																	
5.	2	1	1	1	1	1	8	9	9	12	9	11	13	18	20	20												
6.	1	2	1	2	4	5	6	6	8	7	8	8	12	11	13	10	19	17	20	19	19	20	20					
7.	1	1	1	0	4	2	7	6	7	7	8	9	9	10	14	14	18	17	20	20								
8.	0	0	1	3	3	7	9	10	11	13	12	14	12	14	15	12	15	19	16	20	18	20	19	20	19	19	20	20
9.	0	1	2	2	3	3	7	7	10	8	6	8	18	19	19	20	20											
10.	1	2	4	3	6	4	9	8	10	11	10	13	20	19	19	19	20	20										
11.	1	1	1	1	1	3	3	3	3	4	3	4	17	16	20	20												

* Omitted in computation of results.

TABLE 5

Errors of Group C, Composed of 13 Subjects Who Were Given Two Trials Every Two and Four Days, Alternately

Trials / Subjects	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1.	198	114	246	171	79	39	82	36	28	5	0	0																
2.	214	191	174	140	13	0	1	0	0																			
3.	268	314	186	140	48	8	3	0	0																			
4.	396	381	249	290	53	38	30	4	0	0																		
5.	338	203	217	221	236	143	113	71	118	68	24	15	11	3	0	1	0	0										
6.	454	378	243	298	237	131	169	172	211	136	109	1	0	0														
7.	288	334	272	298	304	153	269	238	143	135	266	205	189	135	44	44	11	30	4	0	1	5	1	0	2	1	0	0
8.	174	202	101	114	115	9	39	42	1	0	0																	
9.	308	160	164	166	159	107	51	7	7	2	0	0																
10.	217	256	222	183	117	66	82	57	1	0	0																	
11.	193	131	68	134	54	20	0	0																				
12.	335	440	321	211	383	225	155	192	193	163	86	114	57	50	14	20	3	0	0									
13.	363	206	227	211	155	63	54	21	25	41	0	0																

TABLE 6

CORRECT RESPONSES OF GROUP **C**, COMPOSED OF 13 SUBJECTS WHO WERE GIVEN TWO TRIALS EVERY TWO AND FOUR DAYS, ALTERNATELY

Trials Subjects	1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1.	2	3	2	7	9	13	10	16	15	19	20	20															
2.	0	1	3	6	7	6	8	7	4	6	9	10	19	20	20												
3.	3	3	3	6	14	16	19	20	20																		
4.	1	0	2	4	14	16	19	17	20	20																	
5.	0	2	2	3	6	7	9	10	10	9	14	16	17	18	20	19	20	20									
6.	2	2	3	6	15	20	19	20	20																		
7.	3	5	5	9	9	17	14	15	19	20	20																
8.	4	2	4	4	5	8	12	16	17	19	20	20															
9.	0	1	2	6	6	11	13	12	19	20	20																
10.	2	4	1	1	1	3	1	1	3	1	3	6	9	10	12	15	16	15	20	19	18	19	20	18	19	20	20
11.	4	3	9	10	10	14	20	20																			
12.	2	1	3	1	2	6	4	6	5	8	9	10	10	14	15	15	19	20	20								
13.	1	1	3	4	5	11	13	14	13	13	20	20															

* Omitted in computation of results.

1 and 2; 3 and 4; etc.), were compared with each other. The same procedure was followed with correct responses but in computing the results Trial 1 was omitted since the correct responses given at that time were due entirely to 'chance'. (Trials 3 and 4, 5 and 6, etc. were compared to determine what happens during the period of practice). To determine the behavior of erroneous and correct responses during the period of rest, the last trial of one day was compared with the first trial on the next practice day, (i.e., Trials 2 and 3; 4 and 5; etc.). The practice days were one day apart for Group A, two days apart for Group B, and two

TABLE 7

PER CENT DECREASE IN ERRORS

	Comparing Trials During Practice		Comparing Trials Separated by Rest Interval				
	Average	Sigma	Average	Sigma	Diff.	Sigma of Diff.	Critical Ratio
Group A	1.2	37.5	26.2	19.9	25.0	12.2	2.05
Group B	24.7	28.5	11.6	32.4	13.1	11.9	1.10
Group C	16.2	31.75	8.4	39.33	7.8	14.37	.54

CORRESPONDING PER CENT INCREASE IN CORRECT RESPONSES

	Comparing Trials During Practice		Comparing Trials Separated by Rest Interval				
	Average	Sigma	Average	Sigma	Diff.	Sigma of Diff.	Critical Ratio
Group A	17.6	14.28	3.99	19.07	13.61	7.42	1.83
Group B	5.1	3.98	16.7	12.32	9.8	3.65	2.68
Group C	5.9	10.21	10.5	9.97	4.6	4.12	1.12

and four days apart, alternately, for Group C. The same method was employed for both errors and correct responses.

The results in Table 7 are the average percentages of decrease in errors and of increase in correct responses during periods of practice and during periods of rest, for all three groups. These values were computed as follows: For each of the three groups the average number of errors and the average number of correct responses were found for each trial. To compare trials during practice, the differences between the averages for successive trials (i.e., 1 and 2; 3 and 4; etc.) were found and these differences stated as percentage of decrease in errors and of increase in

correct responses. The percentages thus derived were averaged for all pairs of trials and these averages are given in Table 7. To determine the behavior of erroneous and of correct responses during the rest periods, the same method was used except that the average errors and correct responses during the last trial on one day were compared with those of the first trial on the succeeding practice day. In computing the averages for correct responses, Trial 1 was omitted for reasons stated above.

The differences between the means of both errors and correct responses during practice and during rest are given in Table 7. Sigmas of the differences between the means and critical ratios are also stated in this table. Several of the comparisons gave actual increases in errors and decreases in correct responses for both practice and rest periods. Such instances were treated negatively and account for the large sigmas in all the tables of results. The results in Table 7 indicate that of the three rest periods investigated, a rest of one day (Group A) is favorable to a decrease in errors, while longer rest periods are actually detrimental. The reliability of the difference between the means for Group A, which rested one day, is greater than for either of the other two groups. Correct responses increase more rapidly during the period of practice when the rest period is for one day only (Group A). The rest period seems more advantageous for increasing correct responses when the interval of rest is continued over two or more days (Groups B and C). However, the reliability of the difference between the means is greatest for Group B. None of the differences between the means is absolutely reliable.

The method employed in obtaining the results given in Table 7 permits a few errors toward the end of the learning period and the few correct responses in the beginning of the learning period to influence the averages too greatly. For instance, Subject 8 (score given in Table 3) is an extreme case. In comparing Trials 21 and 22 the decrease in errors is 100%, although, the numerical decrease is only two. Comparing Trials 23 and 24, the actual decrease is one error but the percentage is again 100. Toward the first of the learning period large decreases in errors

give small percentages of decrease. Correct responses produce the opposite effect, although in this case the percentages are not relatively so large since the total number of possible correct responses in any one interval is never more than twenty, while the possible number of errors is unlimited. In order to eliminate this effect the average errors for each group at each trial were added and the percentage of errors at each trial obtained. To measure the percentage decrease in errors the differences in percentages at each trial were determined. Successive trials were compared and differences computed to obtain the decrease in errors during

TABLE 8

Per Cent Decrease in Errors (Computed from Total Number of Errors)

	Comparing Trials During Practice		Comparing Trials Separated by Rest Interval				
	Average	Sigma	Average	Sigma	Diff.	Sigma of Diff.	Critical Ratio
Group A	1.09	1.03	.49	.44	.60	.33	1.81
Group B	.18	.30	.98	.74	.80	.23	3.48
Group C	.61	.73	.54	.86	.07	.32	.22

Corresponding Per Cent Increase in Correct Responses

	Comparing Trials During Practice		Comparing Trials Separated by Rest Interval				
	Average	Sigma	Average	Sigma	Diff.	Sigma of Diff.	Critical Ratio
Group A	1.14	1.25	—.70	1.36	1.84	.57	3.23
Group B	.17	.18	.26	.39	.09	.12	.75
Group C	.07	.33	.38	.32	.31	.13	2.35

practice. To measure the effect of the rest interval on erroneous responses, the last trial on one day was compared with the first on the next practice day and differences in the percentage of errors computed. Correct responses were treated in the same manner except that Trial 1 was omitted since all correct responses on the first trial were due to 'chance'. An average of all these comparisons is given in Table 8. The change in method of treating the data reverses the results for *error* measurements for Groups A and B, but not in Group C. Results for *correct* re-

sponses are more stable. Group A evidences a more favorable effect from practice than from rest and Groups B and C more favorable effects from rest periods. The results stated in this table justify the conclusion that correct responses increase more rapidly during the period of rest than during the practice period, when rest is longer than one day. The one day interval results in a more rapid increase in correct responses during the period of practice than during rest. The difference between the means in the last instance is reliable.

In the opinion of the writer the values in Table 8 are more representative of the data than are those given in Table 7. However, the varying intervals show no consistent changes, and, since different methods of treating the data give somewhat different results, an analysis of the scores of individual subjects has been made. Table 9 gives the results for each subject, obtained by using the method employed for the group comparisons in Table 8. The trials were compared as before: successive trials were compared to obtain the effects of practice; and the last trial of one day was compared with the first trial on the next practice day to measure the effects of rest. The individual subject's score represents an average of these comparisons, for both errors and correct responses, during practice and during rest periods. The results in Table 9 indicate a wide range of individual differences. Some subjects profit both in their ability to decrease errors and to increase correct responses during the interval of rest. Others decrease their errors and increase their correct responses more during the practice period. For some subjects, the rapid decrease in errors during the practice period is attended by an increase in correct responses during the interval of rest. Still other subjects decrease errors more rapidly during the rest period and increase correct responses more during practice. When the interval between practices is two days or longer, certain subjects actually increase errors and decrease correct responses during practice. Such a condition occurs in only two subjects during the interval of rest. From individual scores in Table 9 the following groups of subjects may be distinguished:

Subjects who find the practice period profitable for the reduction of errors and for the increase in correct responses:

Group A	7 subjects
Group B	1 subject
Group C	3 subjects
Total	11 subjects

Subjects who find the rest period more valuable both for the reduction of errors and for the increase in correct responses:

Group A	5 subjects
Group B	7 subjects
Group C	3 subjects
Total	15 subjects

Subjects who find the practice period more valuable for decreasing errors, but the rest period more valuable for increasing correct responses:

Group A	5 subjects
Group B	1 subject
Group C	3 subjects
Total	9 subjects

Subjects who find the rest period better for decreasing errors and the practice period more profitable for increasing correct responses:

Group A	None
Group B	2 subjects
Group C	4 subjects
Total	6 subjects

When differences in the length of the interval of rest are disregarded, 20 of the subjects decreased their errors during the practice period and 18 decreased theirs during the period of rest, while 3 subjects actually increased their errors during practice; 17 subjects increased their correct responses during the practice period and 21 subjects increased theirs during the period of rest. Three subjects decreased correct responses during practice. Table 10 gives the averages of the individual subject's averages taken from Table 9. When the results obtained by this method of treatment of the date are considered, Group B increased its correct responses and decreased its errors during the period of rest. The differences between the means of the other two groups are too unreliable to warrant any interpretation. The conclusion

TABLE 9

Group A

Subjects	Per Cent Decrease in Errors: Comparing Trials During Practice	Per Cent Decrease in Errors: Comparing Trials Separated by Rest Interval	Per Cent Increase in Correct Responses: Comparing Trials During Practice	Per Cent Increase in Correct Responses: Comparing Trials Separated by Rest Interval
1.	.71*	.53	.09	.42
2.	.03	3.43	.33	.33
3.	6.75	15.95	5.90	11.80
4.	3.47	3.58	.70	1.16
5.	.89*	.68	.29	.34
6.	.79*	.70	.21	.86
7.	1.38*	.11	.58*	.51
8.	1.11	1.13	.02	.73
9.	3.64*	.94	1.27*	.58
10.	2.59*	.22	3.09*	1.33
11.	3.37	11.20	3.47	3.50
12.	3.80*	.60(I)†	2.34*	.86
13.	1.54*	.48	.43	.89
14.	2.09*	.06(I)	.96*	.27
15.	1.84*	.12	.77*	.24
16.	3.13*	.40(I)	2.47	2.80
17.	3.86*	.74(I)	.85*	.14
Group B				
1.	1.00	5.42	.20	1.88
2.	.43(I)	1.51	.41	2.43
3.	.42	1.10	3.27*	.85
4.	.29	1.98	.00	5.40
5.	1.22	2.32	1.93*	1.03
6.	6.34*	3.07	.34(D)‡	1.17
7.	.45	1.20	.26(D)	1.75
8.	.79*	.41	.62*	.05(D)
9.	.69(I)	1.56	.17	2.14
10.	.34(I)	3.75	.13(D)	1.67
11.	.17	1.43	.83	3.27
Group C				
1.	5.36*	1.87	4.92*	.75(D)
2.	2.04*	1.10	.86	2.77
3.	1.36	7.77	3.95	5.70
4.	.32	8.66	.93	6.83
5.	2.34*	.01	.34	1.16
6.	3.13	6.53	5.90*	4.53
7.	1.97	3.45	4.50*	.53
8.	4.38*	1.02	2.45*	2.23
9.	1.57	2.90	3.90*	3.60
10.	.49*	.17	.17	.35
11.	1.70	11.90	5.00	6.00
12.	.51	.89	1.31*	.57
13.	3.98*	.92	2.57*	1.25

* Indicates cases in which errors decrease or correct responses increase more rapidly during practice than during a rest of two and four days alternately.

† Indicates cases in which there is an actual increase in errors.

‡ Indicates cases in which there is an actual decrease in correct responses.

would, therefore, follow that correct responses are increased and errors decreased more rapidly during the period of rest than during the period of practice provided there is an interval of rest of two days.

TABLE 10

AVERAGE OF INDIVIDUAL'S PERCENTAGES OF ERRORS

	Comparing Trials During Practice		Comparing Trials Separated by Rest Interval				
	Average	Sigma	Average	Sigma	Diff.	Sigma of Diff.	Critical Ratio
Group A	2.41	1.06	2.19	2.37	.22	.63	.35
Group B	.84	.86	2.14	.85	1.30	.36	3.57
Group C	2.24	.98	3.34	2.64	.10	.78	.13

AVERAGES OF INDIVIDUAL'S PERCENTAGES OF CORRECT RESPONSES

	Comparing Trials During Practice		Comparing Trials Separated by Rest Interval				
	Average	Sigma	Average	Sigma	Diff.	Sigma of Diff.	Critical Ratio
Group A	1.40	.96	1.57	1.25	.17	.38	.45
Group B	.61	.61	1.96	.78	1.35	.29	4.65
Group C	2.83	1.37	2.67	1.64	.16	.59	.27

CONCLUSIONS

The relative value of the practice period and the period of rest for the reduction of errors and the increase of correct responses varies with individual subjects to such an extent that it would be wise to let each individual determine his own best method of study. Rest periods are advantageous over practice periods for the majority of the subjects used in this experiment, but the difference is slight. Individual differences in the amount of profit and loss from practice and rest are far more evident. Some subjects find the period of practice more favorable for the reduction of errors, but the period of rest more favorable for increasing correct responses. The reverse situation is true for other subjects. It must be remembered, however, that only rational learning was investigated in this study and that the practice periods were probably not long enough to produce fatigue, which might be expected to decrease correct responses and increase errors during the period of practice.

When the data are interpreted from Table 8, it would seem that an interval of rest for two days was most advantageous for decreasing errors and an interval of one day most valuable for increasing correct responses. There appears to be no progressive increase in correct responses or decrease in errors as a result of varying lengths of periods of rest. It is possible that these differences are obscured by individual reactions to the rest and practice periods.

Bibliography

1. Lorge, I. Influence of regularly interpolated time intervals upon subsequent learning. *Teach. Coll. Contrib. Educ.*, 1930, No. 438. Pp. 1-57.
2. Peterson, J. Experiments in rational learning. *Psychol. Rev.*, 1918, **25**, 443–467.
3. Peterson, J. Tentative norms in the rational learning test. *J. Appl. Psychol.*, 1920, **4**, 250–257.
4. Ruch, T. C. Factors influencing the relative economy of massed and distributed practice in learning. *Psychol. Rev.*, 1928, **35**, 19–45.
5. Warden, C. J. The distribution of practice in animal learning. *Comp. Psychol. Monog.*, 1922, **1**, No. 3.

LEARNING TO GENERALIZE

James L. Graham
LEHIGH UNIVERSITY

In generalization problems, the learner's operations are essentially mental; the preliminaries to relational discriminations are usually submerged in chaotic frustrations; and the progress in the earlier stages of the learning can not be directly scored. These difficulties militate against explaining the learning process and skills involved.

This study attempts to provide exercises for generalizing, sufficiently abbreviated to come within reasonable time and motivational limits, and to examine, first, the nature of the tasks, second, some of the skill components involved in successful generalizing, and third, the abilities of college students and the relation of these abilities to certain products of learning. The interest is in presenting a theoretical analysis of this very broad field, based upon the assumption that the generalizing task should be considered as a form of learning. (For a contrasting position, see Maier [23]). The magnitude of this task and the subjective nature of the processes will necessitate the omission of historical material and experimental detail.

The Nature of the Tasks

The tasks could be completely solved by the learning of one or of a few simple relationships which reveal how to get answers for a whole series of problems. They are to be differentiated in whole or in part from related tasks of the following types: those which require the solution of problems by assigning specific answers (24, 33), those involving combination or reorganization of elements or experiences (15, 20, 22), the deduction of useful formulations by the manipulation of accepted statements in accordance with known rules (26), the abstraction of a common uniformity from varying specific elements or experiences (10, 11),

the delayed recall or recognition of previously acquired generalizations or categories of knowledge (38), the application of a known relationship to some new situation (3, 35, 38), and the shift from an accustomed or suggested inhibiting set to a favorable one (19, 21). The general pattern of the problems, materials, directions, sets, and suggestions given as instructions, are described as follows:

1. The instructions required the learning of a method of obtaining answers, from the study of certain specific situations and data. Usually, a formula or verbal description of the derived method of procedure satisfied this goal of learning.

2. An attempt was made to provide both materials for manipulation and an adequate description of the whole problem situation. Typically, this involved a description of a game, materials used, governing rules, methods of play, objectives, and scoring possibilities or comparable equivalents.

3. In order to reduce the number of steps to be learned, and to prevent observational errors which would inhibit the discernment and checking of relationships, specific answers within an arbitrarily selected range, or comparable information regarding the problems were supplied.

4. Specific answers were usually grouped in a more or less systematic arrangement and, typically, involved the presentation of answers in a continuous series beginning with simple forms.

5. In some of the tasks, further arranging, grouping, and other manipulating procedures were suggested, to give the learner an opportunity to observe relationships which were not apparent, or to present them in a different context.

A detailed description of two of these generalizing problems is included in the Appendix to this article (Section A), and a brief description or reference is made to additional generalizations or situations. Some of these are familiar to psychologists in other forms of presentation. The nature of the modification can be inferred from the detailed examples given and the general pattern outlined above. The familiar water-dipping "Ingenuity Test" of the Stanford-Binet Intelligence Scale (36, 37) can be taken for further illustration. The materials furnished are sand

and a graded series of vessels. The types are separated, a graduated series of answers is constructed, and appropriate generalized procedures are required for each. The generalization for the simplest type may be described as manipulating to retain the difference between the large and small vessels. A sample answer series may begin: To measure exactly 3 units, using vessels 4 and 7; fill 4 from 7, and retain 3 in 7.

An accurate understanding of differences in problem-solving tasks is aided by a careful description and analysis of the special conditions imposed by the type of response required, by the kinds and amounts of change necessitated, by the instructions, and by the apparatus used (**30**, p. 387 ff.). The processes involved in conserving, retaining, or applying *old* knowledge, and those involved in discriminating and acquiring *new* or revising the old, may differ enough to justify the development of specialized meanings or new terms, as Maier has attempted to do by restricting learning to situations more like the former, and substituting reasoning for problem solving situations more like the latter (**23**). In any case, the special conditions involved need elucidation, for both the situation and the response aspects.

I turn, then, to five general aspects of learning, to examine them in relation to certain special conditions imposed by my tasks. These aspects are: the learner's starting level, the objective or goal of the learning, the possibilities or limitations of scoring the progress of the learning, the amount of modification necessary to meet the criterion of the learning, and the probable transfer to similar tasks of the sets and skills used.

1. *Special Conditions Related to the Starting Level.* This starting level includes all that the learner brings to the start of the learning, such as: the levels of maturation, intelligence, and familiar knowledge, as well as the skills, habits, attitudes and earlier experience of job satisfaction or dissatisfaction. With reference to the suitability of my tasks and their bearing on where the learner starts, two questions are discussed: (A) What is the relation between the college student's large stock of acquired generalizations and the appropriateness of my tasks to his abili-

ties? (B) What is the relation of his accomplishment to his habituated sets and techniques?

(A) It is difficult to distinguish between new acquisition and recall of previous learning (to tell, for example, whether or not some of the generalizations in the water-dipping problem require new acquisitions or merely the recall of known principles). The criteria I used for determining unfamiliarity included a lack of school instruction upon the generalizations or situations, the student's reports of unfamiliarity, and even a low percentage of successful performance. The finding of suitable but unfamiliar generalizing tasks for college students presents many difficulties, since most relationships which are found in simple settings and which have practical value have been learned. The most suitable problems are low in utilitarian value, and therefore, provide little intrinsic motivation. Furthermore, such tasks are difficult (or give the initial impression of extreme complexity), and hence are unduly disheartening to the learner. Finally tasks for which a complete solution may be derived, in the sense that one or a few relationships will describe how to obtain all correct answers and no others, are mathematical in type, and many individuals have previously developed unfavorable attitudes toward manipulating such symbols and concepts.

(B) The importance of the rôle of habituated sets and attitudinal predispositions is evident. Two examples taken from tasks of a related type may illustrate the influence of accustomed sets. The first is a puzzle employed by Norman Maier (**19**, p. 142) in which six matches are laid on a table (a plane surface) and the subject is required to construct four equilateral triangles out of them, each triangle having a whole match as a side. In this problem, the habituated set to work in a single plane is reinforced by such indirect suggestions as the plane table, the term "side," and by three dimensional manipulative difficulties. Changing the set to working in three dimensions and using the matches as the *edges* of a regular pyramid having a triangular base, afford the solution. The other example is taken from the New York Times' report (**25**) of a problem given to some boys

of the Princeton Country Day School, who referred it to Einstein for solution. The problem required two consecutive odd numbers, the difference between whose squares was –56. It further required a statement as to whether or not the numbers were positive *or* negative. The answer was, " 13 and 15, *and* both." Two special conditions should be mentioned here. One is that the 15^2 be substracted from 13^2 or the larger from the smaller number, a procedure contrary to the accustomed set. The other required the recall of a generalization that the square of either negative or positive numbers is positive, although a set to search for a generalization is not habitual in most people. In tasks of these two types, the special sets, the possible generalizations, and other specific conditions, are so numerous that the successful use of such tasks to develop skills in discerning interrelations is not promising. A more desirable type for such a purpose, I consider to be one that assigns the formulation of a generalization, and, at the same time, reveals the relevant conditions in a fairly comprehensive whole, so that the appropriate direction of the learning is inherent in an understanding of the conditions laid down. A serial form that permits constructing many similar problems, all governed by one or a few principles and similar conditions, offers possibilities As to type, the problem reported by John C. Peterson is a good one (32), although a bit too complex. My modification of this problem is described in detail in the Appendix to this article. The specific answers to the first problem are recognizable by almost anyone who has learned to count by twos; but if a " player " is permitted to choose any number from 1 to 10 inclusive, the basic grouping on each successive pair of plays will usually require a generalization for solution. It should be noted that in this task the formulation of a generalization is assigned, and, furthermore, that the non-habituated set to learn how to obtain answers is enforced by the increasing gradation in complexity.

2. *Special Conditions Related to Scoring Possibilities.* An analysis of the learning of a generalization and of kindred discriminative responses reveals certain limitations and confusion with respect to quantitative scoring. It should be evident that

the scoring of the generalization must be of the all-or-none variety; that is, only full success or complete failure is recognized. I attempted at first to utilize part scores, but was forced by various considerations reluctantly to eliminate them. Part scores (in the form of specific memories and answers, promptings and errors, repetitions and times), when critically examined, failed to yield equal scale units for measuring progress toward learning a generalization since certain changes in set may occur which markedly influence performance and yet which are not adequately represented by part-scores. Part scores which were obtained prior to a shift of set, approach, or technique, were manifestly not comparable to those recorded after such a shift. The attempt to formulate part scores in terms of discernment of the interrelations in the complex situations similarly failed. When the recording of relationships less inclusive than a complete formula was requested, poorer students were often satisfied to report principles which would solve only a few problems or to note an observed uniformity, while better students scorned to record them. When the finding of specific answers for problems somewhat graded for complexity offered scorable items based on lower standards of the learner's progress, poorer students at times utilized illegitimate methods and information to get such answers, while better students rejected their use. Furthermore, in order to construct a problem for the utilization of part scores, it was necessary to stress the learning of elements little related to generalization, or to establish lower standards of mastery. Finally, the utilization of part scores also tended to encourage the learner to adopt inefficient modes of attack (such as memorizing and use of specific information), and fostered the unfavorable set to get specific answers, when the assigned goal was to learn how to obtain answers.

This observation of the limitations of part scores in my own tasks suggests that possibly related studies might profit by a critical examination of the objective scores used with reference to the accurate measurement of the discriminative elements involved in achieving the expected goals. The study by Hull (11) on the evolution of concepts suggests the presence of some of the

difficulties which I found. On the other hand, in tasks where the learning of the problem situation is characterized by suddenness, spontaneity, and completeness, attention is apt to be directed to the final step in the learning because it alone is scorable. In such cases the assumption may be made, perhaps unconsciously, that the scorable end-stage is the total learning. This comment may apply to such a generalization as the following: "Learning standard logical relationships is not a slow, gradual process involving practice-effect or trial-and-error responses but occurs suddenly and gives evidence of permanent mastery." (8, p. 291). If such conclusions may be affected by conditions of all-or-none scoring, they need critical examination to reveal the existence or fallaciousness of an underlying assumption that the scorable final aspects are representative descriptions of the *complete* learning process.

3. *Special Conditions Related to the Goal of the Learning.* The presence or absence of knowledge of improvement affects persistence and job satisfactions. When failure is apparent until complete success is achieved, job satisfactions and motivation are likely to be poor. The value which the learner attaches to the achievement of the goal will affect his motivation, and the immediate usefulness of my goals probably approaches zero, if only their intrinsic worth is considered. The objective of using the exercises as a means of developing general facility in doing such tasks is remote, and the skills desired can not be expected to result from learning one generalization; yet immediacy of use has a favorable advantage for good motivation. The reception of these tasks by the subjects indicated poor motivation, especially for the initial exercise where skills usually are undeveloped. My experience has revealed a need for increasing the subject's satisfactions during that period before success is achieved.

4. *Special Conditions Related to 'Distance' Required to Learn the Task.* Different tasks vary in the amount of modification that must be made to achieve the requisite skills or information. A specific learning process involves modification between two points, that point where the learner starts to learn and the goal

of the learning. Considered as a complete task, the discovery of a major premise has length. I have conceived a generalizing problem as modelled along the lines of the requirements for a Ph.D. dissertation. The ' distance ' in most generalizing exercises is too great for a task of laboratory length or for ordinary motivation. Some short-cutting is necessary. The short-cutting instituted in the tasks which I have used involved the elimination of the collection of ' data,' and the arrangement of the relevant ' data ' so as to facilitate analysis.

5. *Special Conditions Related to Transferability to Other Tasks.* If the learner has developed habitual sets and attitudes which are directionally favorable, and techniques that are efficient, the tasks become relatively smooth running and easy; if, on the other hand, the habitual sets are inhibitory and the skills are inefficient, the tasks become relatively hard and are attended with chaotic frustrations. Assuming that these propositions are valid, the transfer problem becomes one of developing relevant habitual sets and effective skills which are applicable to learning tasks of this type, through the use of suitable exercises and competent instructions. The transferable aspects of generalizing involve responses discriminate in character. They consist of the discernment of influential variables, the discrimination of uniformities in spite of superficial differences, the discernment of how the variables are interrelated and organized and, finally, the formulation of a useful procedure and its verification. The tasks assigned direct attention to these crucial aspects. So analyzed, the learning progress may be expected as a series of discernable " steps " rather than continuous improvement. Initially, my tasks probably give the impression of great difficulty and complexity because of the emphasis on these discriminative aspects alone. However, this artificial isolation of discriminations seems to offer opportunity for perceiving the sets and skills applicable to their learning. The necessity for providing numerous exercises, large amounts of practice, and competent instruction in order to develop transferable sets and skills is, however, not eliminated.

The Skill Components of the Generalizing Task

The skills to be considered are habitual sets, preliminary attitudes, modes of attack, and techniques of manipulation possibly advantageous to making subjective discrimination (28). Initially, the acquisition of such skills is hampered by feelings of chaotic frustration and by the difficulty of measuring discriminative responses. It should be noted that students who succeed, or even prolific makers of generalizations in special fields, fail to show that they can generalize the successful manipulations or discriminative techniques that they use. However, when with several small groups of subjects, an intensive study of as much as six to twelve laboratory hours was devoted to learning such tasks, a marked degree of improvement was conspicuous, as measured by higher percentages of performance in a group. This improvement seems promising for the development of skill components. There was, also, evidence of an increase in job satisfactions in the earlier steps. Since failure is recognized until one begins to formulate the noted uniformities into a useful procedure, sources of satisfaction are small. For the cultivation of such satisfactions, I offer a suggestion derived from the techniques of Edmund Jacobson (12, 13) in teaching "tense" individuals to relax. He first instructs his subjects to contract a given set of muscles maximally, and then to relax them, with a view to teaching them the meaning of muscular hypertension and its reduction. While he stresses the development of *awareness* of the kinesthetic changes, he apparently utilizes the memory of tenseness to develop *satisfactions* from being less tense. Similarly in my tasks, the memories of frustration and of sets that inhibited progress can produce some satisfaction by their very absence when planful manipulations are being used. We may now turn to a consideration of such skill components as habituated sets and techniques.[1]

[1] A greater understanding of the background in which the observations were made may assist in evaluating the material in this section. I first interested myself in generalizing ability, utilizing part scores and six tasks. The performances proved to be poor and part scoring unsatisfactory. Upon modifying the tasks for laboratory exercises, I noted that improvement in ability could be definitely expected with the introduction of attentional factors. When the individual attack was followed by discussion, illustration, and group attack in

1. *The Set to Accept a Generalization as a Goal.* The learner usually expects to get answers, rather than to learn how to get answers. Students often devise non-adaptive goals to which they respond by acquiring more data, checking answers, and manipulating materials beyond all requirements of a clear understanding of the situation.

Let us take, from several possible examples, the familiar bent wire puzzles. The customary task assigned is to learn how to take these puzzles apart and put them together again. It is just as definite an assignment to require the principle of constructing bent wire puzzles so that they can just barely be taken apart and put together again, or to describe the general method of manipulation that will enable the separation of the parts. The accustomed task initiates manipulation; the novel one more likely raises the question, "But what do you want me to do?" Then give them a set of calipers, a finely graduated rule, graduated blade thicknesses, several hardened wire puzzles, and several pieces of soft wire of different diameters and ask specifically, "What are the crucial measurements in the construction of the simple bent nail puzzle?" Students usually do not accept the goal to get the generalization. They want to give you a number such as 1/16 of an inch for the opening, rather than such a generalization as, "The size of the openings that will barely permit separation will be slightly more than one-half the diameter of the wire."

The acceptance of the relational goal will more likely occur if the students have previously learned the desired relationship and only need to recall it or to apply it to a novel situation. In such a case, the learner's effort is directed toward recalling or applying an *old* generalization and away from the set to learn a *new* one. Certainly it is advantageous to accept the experimenter's problem and to cultivate an awareness of the announced goal during all the part activities of the learning.

2. *The Set to Keep all Manipulations Understandable in Terms of How the Parts are Organized in the Whole.* A realization of such an organizing set will be an aid to differentiating the activ-

---

small classes, opportunities for the direct observation of student difficulties, of efficient modes of attack, of the timely introduction of sets, and of the influence of varying directions were afforded. The numerous small sections and need for unfamiliar tasks stimulated the development of additional exercises. Quantitative data were obtained under test conditions from elementary classes, from 523 entering Freshmen, from several classes in experimental psychology, and from individual testing, representing in all about 2,500 papers and 40 different relationships. A brief analysis of certain aspects of the quantitative data will be presented below (*infra*, p. 100). These comments upon skill components and techniques constitute a generalized account based upon experience secured in the extensive experimental program.

ities appropriate to the earlier stages of the generalizing task. The student's past experiences with originals in geometry, with syllogisms, and with algebraic manipulations will have given him a set to start with activities which are applicable, provided some major premise is known. Such a set becomes a barrier to accomplishment when no axiom, relationship, or principle has as yet been learned, since it directs activities toward proceeding accurately from some known principle in accordance with definite rules. Differentiation of the latter set as favorable to the end stages of the task, and an organizational set for the earlier steps will prove advantageous.

Again for our single illustration, let me report observations for the task of finding a short cut method of squaring numbers ending in 5. For this task, I have used the table given below and cardboard pieces to represent units, halves, and quarters, which might be manipulated.

Numbers	Squares
.5	.25
1.5	2.25
2.5	6.25
3.5	12.25
4.5	20.25
5.5	30.25
6.5	42.25
7.5	56.25
8.5	72.25
9.5	90.25

The uniformity of the .25 is readily noted and stimulated the question, "Why?" Also, knowing the relationship of products of the two sides in computing the areas of squares and rectangles, frequently suggests dividing the units of the numbers into the units for the squares. This may suggest noting the uniformity that the quotients are always one more than the divisor. The manipulation of the cardboard halves often suggests utilizing the half units to make whole units. This constructs a rectangle which is always one more than the units on the short side, and always the quarter is left over. After such uniformities have been noted, not before, can the algebraic rules be utilized to aid in securing such a formula as $100X(X+1)+25$, where "X" indicates the number after dropping the first place from the right, or the 5.

3. *The Set to Reduce Complexity by Groupings.* The learner faces frustration often because he seeks to handle too many variables at a time. The complexity of a solid can be markedly reduced if we can deal separately with faces or planes, as can the complexity of areas if we can deal only with linear aspects.

The device of the syllogism keeps the complexity down to two propositions and a relationship, and the derived table reduces it to captions, stubs, and a relationship. The set to combine items in order to make a similar reduction in complexity is an important manipulative procedure, and in such activities the student can derive satisfaction. One method is to substitute a common name for many discrete items.

For example, in the problem for obtaining a short-cut method for squaring numbers ending in 5, naming the units in the table under numbers, "the short side of the reconstructed rectangle" (the 'X' of the formula) and the series of quotients 2, 3, 4, 5, etc., "the long side of the rectangle" ('X+1' of the formula) reduces the task to discriminable dimensions and increases the probability of formulating a generalization.

I have increased by several hundred percent the incidences of success for several tasks in supposedly comparable groups, when I incidentally suggested such combining. This occurred for different revisions of John C. Peterson's Bead Problem, described in modified form in the Appendix, by a mere reference in the description of the task to certain possible choices as the lowest and the highest choices permitted. This set-to-combine reduces the difficulty of discriminating and, in addition, may fortify other favorable organizational sets.

I have given Dr. Joseph Peterson's Rational Learning problem (**29**) to over 200 adults in the 10 letter form and have analyzed the results for evidence that they had found two favorable methods of organization sufficiently early in the task to eliminate efficiently errors of the logical and perseverative type. In almost all cases, they had failed to leave clear evidence of having utilized both types of favorable organization early in the learning. Modifying the instructions, I suggested the desirability of grouping and stated that a long license plate number of 6 or 9 digits on a moving car could be more easily perceived and remembered if these were considered as 2 or 3 groups of three place numbers each. In analyzing the data of the group to whom was given the grouping set, it was found that both organizational procedures which enabled the elimination of the logical and perseverative errors were adopted at a sufficiently early stage to be definitely helpful. An additional principle was reported by a number of the group receiving the combining instructions, but reported by none of the other group, namely, discarding the digit-letter substitution instructions and learning the problem as a straight position-order task. Furthermore, the level of difficulty was so greatly reduced that the median scores for the 8 and 10 letter forms of this problem were in excess of the 90th percentiles of the non-directed groups for time, repetitions, and all types of errors. Also, the median scores for the much more difficult 16 letter form exceeded the 75th percentile of the easier 10 letter form of the non-directed group. The poorest student in the 16 letter form equalled the 40th percentile

of the 10 letter form group. It should be noted that all subjects taking the 16 letter form had previously had the 10 letter form as a fore-exercise.

4. *The Set to Use Combining and Grouping Techniques as Aids in Discriminating Relationships.* The familiar techniques of the syllogism and derived table are customarily used by instructors to aid them in presenting relationships or by students as an aid in retaining them. By some modification in sets and procedures, such techniques may be employed to facilitate the discovery of relationships. The usefulness of the derived table has become evident in learning several tasks, both in connection with guidance toward desirable sets and procedures, and in test construction. When employed as a device with appropriate sets, it may be used both to aid in discerning uniformities and to assist in the discernment of relationships. The arrangement of the data into this tabular mould directs attention to the discrimination of those variables that are likely to prove influential, since the latter are likely to be selected as stubs for the rows and as captions for the columns. It also directs attention to a systematic arrangement of the data into appropriate cells. With the set to reduce complexity by combining the discrete answers into common categories based on observed uniformities, the systematic arrangement of the data facilities the discernment of such uniformities. With the set to discover relationships between important variables, attention can be directed to the boxes intersecting row and column headings, which in turn means that the learner is working at a level that should admit hope of success. The derived table also enables the subject to utilize attentional factors and principles of learning and of perceptual organization (14, 16, 39, 40). Some of these principles which, with slight modification, are found to be advantageous in the tabular arrangement may be noted:

a. Elements that are in close proximity are likely to be united organizationally. The tabular form gives a visual unity, compactness, and proximity seldom found in other devices.

b. Relationships are more easily discerned in a simple setting than in a complex one. In presenting relationships, comprehension is made easy by a simple situation. In learning relationships,

however, the simple setting gives answers so easily that the relationship is unsought. It is only when answers become difficult to get that a felt need for a relationship develops. The derived table may then present a situation favoring return to the simpler forms. These are usually given in the upper right hand corner of the table—a position relatively advantageous for attention-getting.

A single example of this advantage may be illustrated with the Disc Transfer Problem (**31**, p. 31). The correct manipulation is expected for a pile of three discs. There is no need to use simpler problems as fore-exercises. Yet a marked rise in percentages of successful generalizations was noted when answers for 1 and 2 discs in a pile were included in the table. Having once discerned that the fundamental uniformity is the successive doubling, the necessary qualification becomes simple only when it is noted that for 1 disc, one moves directly to the appropriate station. The formula 2^n-1 for the series of answers for the number of discs beginning with 1, namely, 1, 3, 7, 15, 31, 63, 127, etc., becomes discernible, where "n" represents both the power and the number of discs in the pile. Yet a successful generalization is difficult to learn without the return to the simplest form.

c. The favorable advantage of the principles of continuity and completeness of data has been well demonstrated in organizing figures presented tachistoscopically. These principles are customarily utilized in the tabular arrangement for a range that permits ready discrimination.

d. The tabular arrangement usually offers a systematization that favors utilizing the principle of inclusiveness and facilitates the perceiving of recurring cycles, rhythms, or various pattern groupings.

e. Another advantage accruing from a table well-filled with answers is that it affords a means of ready, prompt, and accurate checking of uniformities and hypotheses. In several tasks the use of a table has been observed to forestall too narrow and limited generalizations, due to the ease of checking made possible by its use.

f. The tabular device has also enabled the learner who has discerned a similarity, but one which needed revision or qualification to give answers, to hang on to the unworkable principle and to accept the set to revise it in order to make it workable rather than to discard it.

I do not offer the derived table as a royal road to the discovery of relationships, but as one technique which observation of several

tasks has shown to be successful when applied with appropriate sets.

5. *The Set to Try Out Alternative Possibilities.* The tendency to remain in habituated directional ruts has been discussed. It can be avoided by a set to attend to the systematic surveying of possible alternative approaches. Maier has shown that for tasks where inhibitory sets run counter to success, the value of warning the student to avoid habitual directions has increased the number of problems solved from 20 to 40 percent (23, p. 372).

6. *Sets and Techniques to Control Emotional, Affective, and Attitudinal Predispositions.* Most subjects need to recognize the magnitude of emotional handicaps and to adopt sets and techniques to lessen them. In order to overcome them, I suggested to my more intensive study groups the use of the sets and techniques employed by the scientists, jurists, and logicians. For these groups, less evidence of inferiority feelings, and greater disinterestedness, was noted for later than for earlier tasks. The techniques of the scientist might be described briefly as controlled observation, objectivity in the measurement of his observations, quantitative description or measurement, and the use of mathematical, mechanical, and statistical concepts to facilitate grouping and analysis. To these techniques he brings the pose of a disinterested observer seeking to enforce the coöperation of nature by his controlled questionings and observations. The jurist's techniques seek the admission of all relevant evidence that each side brings to his attention and the suspension of judgment until the facts are in, and the similar pose of weighing the facts disinterestedly and impartially without special interest or favor. The logician's technique involves adherence to the rigorous rules of the game in his manipulations of language and symbols, and the same pose of impersonal, impartial disinterestedness in outcomes. I use the term 'pose' advisedly, since poses can be cultivated which lessen the handicaps of felt inferiority and frustration. I include systematic techniques, since it is highly desirable that generalizing be regarded as a form of learning rather than as a vague 'higher' process.

7. *The Influence of 'Time Pressure' Upon Goal Achievement.* It was observed on a few occasions that higher percentages of

successful performance occurred in a group of subjects who were allowed approximately half the time allotted to a presumably comparable group for work on the same tasks. Since the percentages of success were small in both cases and since the observation could have been explained on other grounds, the problem was further studied in a related setting. No verbal formulation of the generalization was required and the checking of specific answers, yielding part scores, was assigned. The question raised was whether or not an impossibly long assignment would tend to enforce the generalizing set and an effort to educe short-cuts, where a more reasonable time limit would enforce the more habitual but less effective sets and methods. Specific answers were required in the following tasks and no verbal formulation of a generalized procedure was assigned.

Tasks were given to two groups of approximately 40 students each. One task was composed of 330 numbers, and two forms were made. The numbers were arranged in numerical order in eleven columns. One form gave numbers ending in 1 starting from 1, the other, those ending in 7. The time limits were 3 and 9 minutes. Both forms were given to each group. One group was given the 3 minute form first, followed by the 9; the other group received the 9 minute form first. Instructions were: "Cross out all non-prime numbers. Your score will be the number crossed out minus three times the wrongs. Use any method, but do not forget that your object is to make a maximum score in the time allowed." I arbitrarily eliminated all records where scores were less than 50 as probably not utilizing short-cuts. I analyzed only scores that met this criterion for the separate forms. Of the group receiving the 3 minute form first, 57% succeeded, to 54% for the group receiving the 9 minute form first, a gain for the time-pressure group of 3%. On the internal evidence of a generalization, the percentage for the group receiving the 3 minute form first is 50; and for the 9 minute form first, 36, a gain of 14% in favor of the 'time-pressure' group.

In order to compare this gain roughly with that dependent upon the giving or withholding knowledge of short-cuts, this task was followed by a second task, which consisted of 330 even numbers of three and four digits each, arranged in a haphazard order, with instructions to cross out numbers divisible by 4 or by 6. They were scored similarly. Employing the same method of comparison, 49% of the group with knowledge of short-cuts and 30% of the group from whom this knowledge was withheld succeeded in meeting the criterion of scores of 50, a gain of 19% attributable to the suggestion to look for short-cuts. The analysis for internal evidence that a generalization was formed gave 46% and 30% respectively, a gain of 16% in favor of the group having knowledge of the existence of possible short-cuts.

These results suggested that both a set to look for generalizations and a feeling of being under the pressure of time, may enforce the set toward relational organization and tend to in-

crease an adoption of new methods which may prove more efficient than habitual ones.

The skill components which have been suggested in this section are in large part attentional factors, attitudinal predispositions, and procedures of manipulation which may be acquired by an individual. Their development is suggested as a way of increasing planful manipulation, of reducing inhibiting sets and procedures, and of aiding the cultivation of job satisfactions in those stages where only failure can be scored over perhaps a long period, even in the cases where successful generalizations are eventually accomplished.

Generalizing Abilities of College Students and Their Correlation With Other Learning Abilities

The incidences of successful performance for unfamiliar generalizing tasks of the type described were low for the groups of untrained Lehigh lower-classmen. Taking fifteen tasks,[2] with no definite instructions concerning economical approaches or techniques of arrangement, re-grouping, or manipulation, the median percentage of successful performance was 11, and the mean percentage was 18.6. In every case, rather narrow limits of time were imposed, since few students working in a group will persist in attacking such problems, continuously recognizing failure, without creating unfavorable group attitudes. No student was permitted to work on a related series of problems longer than 90 minutes. Many were limited to 30 minutes or were given permission to return the task for an easier one, if willing to admit defeat. For three tasks,[3] where additional suggestions were included in the instructions to arrange systematically, to re-group, or to manipulate, the median number of successful performances was increased to 33% and the mean to 24.4%. The number of

[2] See Appendix to this article, problems 1–9 inclusive and 15, 18, 20, 21, 23, and 24. Certain relationship assigned which might possibly be recalled as in portions of 8 and 21 were omitted, and even some of the easier presentations of 10, 14, 17, 19, 26, and 27 were likewise excluded in an effort to secure a rather homogeneous group of tasks and instructions.

[3] See problems 2, 13, and 23. For illustration of the meaning of "additional suggestions" note the final paragraphs in the detailed descriptions of the problems listed in Part I of the Appendix. These are problems 23 and 13.

subjects tested in the above 18 tasks ranged from 590 to 63, for the various tasks. When training was instituted, amounting to from 6 to 12 laboratory hours, and consisting of illustrating the generalized sets and techniques noted in the previous section, and the assignment of additional exercises, the incidence of successes was increased several hundred percent, for several different tasks.

The ability to learn the fifteen tasks without such special suggestions was analyzed for the successful groups on each task, with reference to relationship to the students' general scholastic average (making use of the University's quality point system), and to percentile ratings on the American Council on Education Psychological Examination. The ability to generalize was somewhat positively related both to higher scholarship and to the intelligence test scores. Without presenting the data in detail, the following summarization will indicate something of the degree of the relationships. On the average, about 59 percent of the successful group surpassed the university average scholarship median and about 15 percent surpassed the 90th percentile. None of the students in our 'successful' groups was found to be in the category of students making the lowest passing grade. The relationship of success to scores on the American Council on Education test was higher. About 73 percent surpassed the University median, and about 30 percent surpassed the 90th percentile, while only about 10 percent of the successful group were in the lower quartile. Furthermore, the prediction can be quite confidently made that if a student is not above the lower quartile in either general scholastic average or American Council on Education test score, he will not succeed in learning generalizations of this level of difficulty without additional guidance and suggestions.

Conclusion [4]

This study considers generalizing as a form of learning, examines some special conditions imposed by the type of learning,

[4] At the recent Columbus meeting of the American Psychological Association, Dr. John C. Peterson told me that his brother Joseph had undertaken the preparation of an article on *Insight* and had requested his full notes on the generalizing problem he had published in 1920 (**32**) to support this article. I regard it as a matter of deep regret that death left this desire unconsummated and this contribution unpublished.

develops a general pattern for the construction of exercises to aid in developing skills in generalizing, and illustrates this pattern with detailed examples and references to other tasks employed. It stresses that attention be given to the earlier stages of the learning, to the development of job satisfactions, habitual sets, and techniques appropriate to early stages, and to reducing the length of the tasks to include only the crucial aspects of the generalizing problem. In the support of the observations, inferences, and conclusions given, the experimental evidence is utilized largely in an illustrative manner, due to the wide general scope of the material covered, to the lack of ability of obtaining parts scores for analyzing the skill components, to the difficulty experienced by subjects in reporting their planful manipulations, sets, techniques, or processes for the illumination of their activities, and to the fact that the crucial aspects of the learning are so essentially subjective mental discriminations. The details of experimental findings, therefore, are largely omitted. The limitation of scoring, and the subjective nature of the data available for analysis raise the very fundamental problem of developing new modes of investigating the early stages of the learning in complex settings. In two former studies (6, 7), I have noted the relationship between illusions in very complex settings and the arrest and fixation of the learning process at the early unskilled stages. An experimental attack upon the reasons and conditions which effect such fixations at low learning levels might well throw light upon these obscure and submerged phases.

One quite apparent difficulty is that of selecting tasks for generalizing exercises at the college level. This difficulty is not so great in the lower grades and in animal problems where the stock of simple relationships in simple situations is small but is greatly increased at the post graduate level where generalizing becomes an educational requirement. Adapting the exercises to the student's degree of attainment becomes increasingly difficult as the stock of acquired generalizations increases. Group projects that distribute the drudgery of complete tasks may perhaps abolish the need for short-cutting, increase the motivation to carry through the tasks to completion, reduce the sense of in-

feriority and emotional disruption, and admit of more competent guidance and instruction in connection with such learning.

The advantages of training for skill in generalizing far surpass the hope that important new generalizations will thus be increased. The by-products may be even more valuable in developing attitudes toward the nature and origin of principles, and in releasing learners from the binding effects of inadequate solutions, predispositions, skills, and methods that hold them to low performance levels. Any education is one sided that merely conserves acquired habits, and attitudes. Skills also are needed that equip the individual to adapt successfully to changing cultural demands.

Bibliography

1. Andrews, E. E. An excursion in numbers. *Atlantic Monthly*, 1934, **154**, 459–466.
2. Ball, W. W. R. *Mathematical Recreations and Essays*. London: Macmillan, 1922. Pp. 366.
3. Ballard, P. B. Noegenesis. *J. Gen. Psychol.*, 1929, **2**, 433–443.
4. Boulton, C. L. Nim: A game with a complete mathematical theory. *Annals of Math.*, 1901, **2**, 35–39.
5. Dantzig, T. *Number: The Language of Science*. New York: Macmillan, 1930. Pp. 260.
6. Graham, J. L. Illusory trends in the observation of bar graphs. *J. Exper. Psychol.*, 1937, **20**, 597–608.
7. ——— Illusory trends in the observation of volumes. *J. Exper. Psychol.*, 1937, **20**, 609–619.
8. Hartman, G. W., and Schooley, M., The rôle of insight in the learning of logical relations. *Amer. J. Psychol.*, 1937, **49**, 287–292.
9. Haught, B. F. The interrelation of some higher learning processes. *Psychol. Monog.*, 1922, **30**, No. 139. Pp. 71.
10. Heidbreder, E. An experimental study of thinking. *Arch. Psychol.*, 1924, **11**, No. 73. Pp. 175.
11. Hull, C. L. Quantitative aspects of the evolution of concepts. *Psychol. Monog.*, 1920, **28**, No. 123. Pp. 86.
12. Jacobson, E. *Progressive Relaxation*. Chicago: Univ. Chicago Press, 1929. Pp. 429.
13. ——— *You Must Relax*. New York: McGraw-Hill Co., 1934. Pp. 201.
14. Koffka, K. *Principles of Gestalt Psychology*. New York: Harcourt Brace, 1935. Pp. 625.
15. Köhler, W. *Mentality of Apes*. New York: Harcourt Brace and Co., 1925. Pp. 342.
16. ——— *Gestalt Psychology*. New York: Liveright, 1929. Pp. 403.
17. Lehmer, D. N. Factor table. *Carnegie Inst. Wash.*, 1909, No. 105.
18. MacMahon, P. A. *New Mathematical Pastimes*. Cambridge: The University Press, 1930. Pp. 116.
19. Maier, N. R. F. Reasoning in humans: I. On direction. *J. Comp. Psychol.*, 1930, **10**, 115–143.

20. ——— Reasoning and learning. *Psychol. Rev.*, 1931, **38**, 332–346.
21. ——— An aspect of human reasoning. *Brit. J. Psychol.*, 1933, **24**, 144–155.
22. ——— Reasoning in children. *J. Comp. Psychol.*, 1936, **21**, 357–366.
23. ——— Reasoning in rats and human beings. *Psychol. Rev.*, 1937, **44**, 365–378.
24. Mather, J. E., and Kline, L. W. The psychology of solving puzzle problems. *Ped. Sem.*, 1922, **29**, 269–282.
25. *New York Times*, 1937, June 12 and 17, Editorial pages.
26. Nunn, T. P. *The Teaching of Algebra.* New York: Longmans, 1931. Pp. 616.
27. Peterson, G. M. An empirical study of the ability to generalize. *J. Gen. Psychol.*, 1932, **6**, 90–114.
28. Peterson, J. The effect of attitude on reproduction. *J. Educ. Psychol.*, 1916, **7**, 523–532.
29. ——— Experiments in rational learning. *Psychol. Rev.*, 1918, **25**, 441–467.
30. ——— Intelligence and learning. *Psychol. Rev.*, 1922, **29**, 366–389.
31. ———, and Lanier, L. H. Studies in the comparative abilities of whites and negroes. *Ment. Meas. Monog.*, 1929, **5**. Pp. 156.
32. Peterson, J. C. The higher mental processes in learning. *Psychol. Monog.*, 1920, **27**, 129.
33. Ruger, H. A. The psychology of efficiency. *Arch. Psychol.*, 1910, **19**, No. 15. Pp. 88.
34. Smail, L. L. *College Algebra.* New York: McGraw-Hill, 1931. Pp. 450.
35. Spearman, C. *The Nature of Intelligence and the Principles of Cognition.* London: Macmillan, 1923. Pp. 358.
36. Terman, L. M. *The Measurement of Intelligence.* New York: Houghton Mifflin, 1916. Pp. 362.
37. ———, and Merrill, M. A. *Measuring Intelligence.* New York: Houghton Mifflin, 1937. Pp. 461.
38. Touton, F. C. Solving geometric originals. *T. C. Contrib. to Educ.*, 1924, No. 146. Pp. 115.
39. Wertheimer, M. Untersuchungen zur Lehre von der Gestalt. *Psychol. Forschung.*, 1923, **4**, 301–350.
40. ——— Zu dem Problem der Unterscheidung von Einzelinhalt und Teil. *Zeitsch. f. Psychol. u. Physiol. d. Sinnesorgane*, 1933, **129**, 353–357.

APPENDIX A

A Detailed Description of Two Problems

I. *A Modification of John C. Peterson's Bead Problem* (32).

I. *The Conditions of the Game and the Rules Governing it.* The materials of the game can be any sort of counters (beads, beans, chips, coins). The number of counters placed upon the table at the opening of any game may be any number agreed upon or arbitrarily determined. The two opponents play alternately. A play is made by removing one or more counters, limited to the number permitted, from the original number or from those remaining after preceding plays. The number of counters which may be removed at any play is any arbitrarily determined number. The object of the game is to remove the last counter or to make the final play.

2. *Objective.* (Note that a generalization is assigned.) Describe a method or methods of obtaining " safe combinations," that is, combinations which will enable the player who leaves them to win eventually, even though both players play expertly thereafter.

3. *Safe Combinations.* A ' safe combination ' is any number of counters which under the rules of the game will not permit the opponent to win. It can be shown that if player "A" can leave a safe combination for " B," then " B " cannot in turn leave any safe combination for " A," but that "A" on the following play can always leave " B " another safe combination. Hence, as the counters are reduced by successive plays, "A" may always return a safe combination from any safe combination which on the previous play he has left " B," and " B " can never leave such a combination. Thus proceeding, "A" always leaving " B " safe combinations and " B " never able to do so, "A" will eventually win by taking the last counter or making the last play. The generalization problem assigned requires learning how to determine these safe combinations. A list of the safe combinations for a limited selection of choices and for a limited range of open-

ing numbers is given below, in order to make the determination of such safe combinations unnecessary and to make the checking of hypotheses easier.

4. *Interpretation of the Tables.* In the column headings are given all numbers of counters that the players are permitted to remove at a given play. In the column below each heading are given all safe combinations for these choices within the limited range of the smallest nine arranged in consecutive order from the lowest opening number of counters. In reading the column headings, or choices permitted, the hyphen indicates that all numbers inclusive of the first and last may be removed; and the comma indicates that only the numbers indicated are permitted. For example, ' 1–6 ' means that 1, 2, 3, 4, 5, or 6 counters may be removed at any play, but ' 1, 6 ' means that only 1 or 6 may be removed.

TABLE 1

SHOWING 'SAFE COMBINATIONS' FOR THE RANGES OF CHOICES INDICATED IN THE COLUMN HEADINGS (CONTINUOUS SERIES)

Permitted Choices								
1	1–2	1–3	1–4	1–5	1–6	1–7	1–8	1–9
2	3	4	5	6	7	8	9	10
4	6	8	10	12	14	16	18	20
6	9	12	15	18	21	24	27	30
8	12	16	20	24	28	32	36	40
10	15	20	25	30	35	40	45	50
12	18	24	30	36	42	48	54	60
14	21	28	35	42	49	56	63	70
16	24	32	40	48	56	64	72	80
18	27	36	45	54	63	72	81	90

Permitted Choices								
2–3	2–4	2–5	2–6	2–7	3–4	3–5	3–6	4–6
5	6	7	8	9	7	8	9	10
6	7	8	9	10	8	9	10	11
10	12	14	16	18	9	10	11	12
11	13	15	17	19	14	16	18	13
15	18	21	24	27	15	17	19	20
16	19	22	25	28	16	18	20	21
20	24	28	32	36	21	24	27	22
21	25	29	33	37	22	25	28	23
25	30	35	40	45	23	26	29	30

All the answers in the first block and some of those in the second may be secured by the learning of a simple relationship.

The additional answers in the second block may be determined by a supplementary relationship, or may be stated as a qualification to the fundamental principle. Describe the method that will enable you to determine these safe combinations. These relationships should be sufficient to enable determination of the safe combinations for all those of similar continuous series, that is, all situations in which any number may be removed at any play between and inclusive of the highest and lowest choices permitted.

The next two blocks of answers might be designated as a discontinuance series, in that the intermediate numbers are not permitted. Either an additional principle is needed for them or a still broader or more inclusive statement of the fundamental principle is required. Note that a statement of principles that is adequate, but which is not the most inclusive statement of the

TABLE 2

DISCONTINUOUS SERIES: PERMITTED CHOICES ARE INDICATED IN THE COLUMN HEADINGS, AND THE CORRESPONDING 'SAFE COMBINATIONS' ARE SHOWN IN THE COLUMNS BELOW

Permitted Choices							
1, 3	1, 5	1, 7	1, 9	1, 4	1, 6	1, 8	1, 10
2	2	2	2	2	2	2	2
4	4	4	4	5	4	4	4
6	6	6	6	7	7	6	6
8	8	8	8	10	9	9	8
10	10	10	10	12	11	11	11
12	12	12	12	15	14	13	13
14	14	14	14	17	16	15	15
16	16	16	16	20	18	18	17
18	18	18	18	22	21	20	19

Permitted Choices									
2, 6	2, 10	3, 9	3, 15	4, 12	2, 4	2, 5	3, 7	4, 9	3, 5
4	4	6	6	8	6	4	6	8	8
5	5	7	7	9	7	7	10	13	9
8	8	8	8	10	12	8	11	14	10
9	9	12	12	11	13	11	12	15	16
12	12	13	13	16	18	14	16	16	17
13	13	14	14	17	*19*	15	*20*	21	*18*
16	16	18	18	18	(4)	*18*	(7)	*26*	(6)
17	17	19	19	19	(5)	(5)	(8)	(9)	(7)
20	20	20	20	24	(10)	(12)	(17)	(10)	(14)
21	21	24	24	25	(11)	(19)	(18)	(11)	(15)

fundamental relationship, will without qualification give the losing combinations inclosed in parentheses. See if the relationships can be so qualified or stated that these losing combinations will be eliminated.

4. *Solution.* The fundamental relationship pertains to understanding the organizational possibilities for a single pair of alternate plays. Such safe combinations may be expressed as the lowest plus the highest permitted choices, or, for an added principle in the discontinuance series, as twice the lowest choice. Multiples of these numbers will also be safe combinations, since by grouping one may, also, always divide the original number of counters into groupings of the simpler problems. Due to the fact that the final play is also a winner, any combination that leaves less than the lowest permissable choice becomes a safe combination. Hence, a supplementary principle that 1, 2, 3, 4, 5, up to the lowest choice permitted minus 1 should be added. Qualifications of the fundamental principle of the lowest plus the highest choices permitted are not required in either series. However, multiples of twice the low are nullified by the intermediate choices in the continuous series. In the discontinuous series, the multiples of twice the low are only applicable if they are less than the highest choice. The supplementary principle of 1, 2, 3, etc., up to the lowest choice minus 1, when applied to twice the low multiples is similarly modified by the high choice.

This problem was added to the group of "additional suggestions" to take hold to manipulate planfully, by suggesting a survey of the "safe combinations" and all possible manipulations involving opening numbers not greater than the highest that could possibly be reduced by a single pair of alternate plays for B to win.

II. *Problems Involving the Use of Factor Stencils* (17).

Objectives. Problem 1: Describe a method of applying short factor stencils to a long table to obtain all numbers evenly divisible by the specific factor of the key stencil. State the relationship in such general terms that the derived procedure is sufficient for

any prime factor stencil. Problem 2: Describe a method of utilizing prime factor stencil keys to obtain non-prime factor stencil keys without additional stencil cutting. Note that the prime stencil keys may require added length to give all answers. The following supplementary principle was added: Describe in general terms the minimum length of the key required to give all correct answers. Problem 3: Describe a method which will tell what stencils, other than prime stencils, must be cut to obtain stencil keys for all possible factors.

1. *Materials Provided.* 1. A set of perforated stencil keys for the earlier prime number factors, namely, 2, 3, 5, 7, and 11. 2. A tabular arrangement to identify consecutive numbers up to 510 for use with the short prime stencils. 3. Directions for the application of the short stencil keys for the top of the table, namely, for numbers 1–99 inclusive. 4. Added suggestions for manipulating in a different context for the two later problems.

2. *Interpretational Comments.* 1. A " factor stencil " is a stencil showing perforations, which, when properly applied to the tabular arrangement for identifying consecutive numbers, will indicate those numbers and only those numbers that are evenly divisible by the factor for that stencil. 2. A " prime factor " is a prime number, that is, a number which can only be divided by itself and 1. 3. A " relationship " is a description of a connection, which, in this case bears upon the question of how to proceed to get answers. 4. The method of using the table for identifying consecutive numbers is as follows: The numbers heading the column represent the units of any number. Note that these numbers also appear on the stencil keys. The numbers heading the rows are consecutive multiples of ten. Any number up to 510 can be identified by noting the position of lines intersecting row and column headings and adding the units of the column headings to the number of the corresponding row. Thus the number 39, the sum of 30 and 9, is found at the intersection of lines for the column headed 9 and the row headed 30. Similarly 130 is found at the intersection of the row headed 130 and the column headed 0.

3. *Manipulative Procedures and Suggestive Arrangements.* (a) For Problem 1: Place each stencil provided upon the table so that the zero of the top row of the stencil is adjusted upon the zero of the table. Note that the perforations give correct answers for the factor used, up to 99. Learn to adjust the stencil upon the various parts of the table so that correct answers and only correct answers are given. Seek to formulate in general terms instructions for adjusting these stencils so that they will be equally applicable for the adjustment of any of the five stencils provided. Note that for prime stencils higher than 11, longer stencils must be utilized.

(b) For Problem 2: If the stencils and table are properly adjusted for length, the following additional factor stencils may be made, using only the five stencils provided: 6, 10, 14, 15, 21, 22, 30, 33, 35, 42, 55, 66, 70, 77, 105, 110, 154, 165, 210, 231, 330, 385, 462, 770, 1155, 2310. Learn to construct such stencils and to formulate the derived method so that one set of instructions will apply to the various factors given above. Starting with the lowest factor in the list, search for a method of using the prime stencils to obtain a stencil for the new factor, then try it for the next factor higher, etc. Test your method to determine whether or not it is stated broadly enough to apply it to the factors for 30, 42, 66, 70, and higher numbered factors. If you fail to discover a method of constructing such stencils, it is suggested that you utilize the different context and procedure given in the final paragraph of Problem 3.

(c) For Problem 3: In order to provide factor stencil keys for all factors, 2–100 inclusive, additional prime factor stencils must be cut, namely, 13, 17, 19, 23, 29, 31, 37, 41, 43, 53, 59, 61, 71, 73, 79, 83, 89, and 97, but only the following non-prime factor stencils, namely, 4, 8, 9, 16, 25, 27, 32, 49, 64, and 81 need be cut. Learn how to determine just what non-prime stencils must be cut and to formulate a generalization to describe the derived procedure.

The following context and procedure may assist you in seeking the desired relationship. Place in a row the consecutive numbers

beginning with 1, as in the arrangement below. Then write below each number the lowest prime factor which will factor that number, then below that, the lowest prime factor that will factor the quotient obtained, etc., until the numbers may be represented as the products of the prime factors. Study the arrangement for uniformities existing in the correct answers.

1,	2,	3,	*4*,	5,	6,	7,	*8*,	*9*,	10,	11,	12,	13,	14,	15,	*16*,	17,	18,	19,	20,	21,	22,	23,	24,	*25*,	etc.
			2		2		2	3	2		2		2	3	2		2		2	3	2		2	5	
			2		3		2	3	5		2		7	5	2		3		2	7	11		2	5	
							2				3				2		3		5				2		
															2								3		

APPENDIX B

A Classified List of Exercises for Developing Generalizing Ability

The difficulty of finding suitable tasks for the college and superior adult level has been commented upon. I am, therefore, including a list of the tasks which I have used, or which, in a few cases, are similar to those used. The tasks vary in inherent difficulty and complexity, and with method of presentation.

I. *Tasks involving summation series:* The derivation of a *method* of computing answers to the following problems is required, and appropriate illustrative answers for a limited range of each series is stated.

1. Using a series of equilateral triangles with the area uniformly dotted and the triangle built up with 1,2; 1,2,3; 1,2,3,4; etc., dots in a row, a method of computing the number of dots in the triangle from knowledge of the number of dots on a side is required (5, pp. 41–43).

2. Using a series of squares with 2,3,4,5, etc., dots in rows and columns, each square being divided into two right triangles, a larger one employing the greatest diagonal, the smaller one, the adjacent exterior diagonal, a method of computing the number of dots in the large and, also, in the small triangle, from knowledge of the number of dots on the side of the square, is required (5, pp. 41–43).

3. Using the faces of regular solids for figures, the method of computing the total dots in the sides of the figure, from knowledge of the number of dots on any side is required. Similar tasks are assigned, involving the edges of the regular solids, and areal tasks, involving one or all faces of the regular solids.

4. Using series of pyramids having equilateral triangle or square bases, and marbles in the place of dots, tasks similar to the above are given, requiring a method of computing the number of marbles in a pyramid from knowledge of the number of marbles on the edge of the pyramid.

5. Using the construct of Pascal's triangle, and the terms of

the series in any column, the subject is asked to develop a method of getting answers for the various columns' series and a general formula for all series from knowledge of the number of the term. A supplementary principle is required for a method of constructing this triangle (**34**, p. 180).

6. Using a series of squares ruled off into one centimeter units of 1, 2, 3, 4, etc., centimeters in a row or column, a method of computing the number of different squares possible from knowledge of the number of centimeters in any row is required. Similar tasks are set for the number of different rectangles, remembering that a square is always a rectangle but that a rectangle is not always a square; and for a cube, to determine the number of different cubes in a cube, and the number of rectangular solids in a given cube.

II. *Tasks involving difference series:*

7. A series of tasks is assigned, utilizing interpolation tables and supplementary tables developing the method of constant differences. The tables give only integer values, and a method of interpolation to secure accurate answers for decimal values is required.

III. *Tasks involving product series:* In general, from knowledge of a number in an arithmetical series and appropriate answers for a limited range, learning a method of computing answers is required.

8. Using Atkinson's Ingenuity Test situation, the method of computing the number of possible combinations from knowledge of the number of letters in any row is required. A systematic method of arranging answers to eliminate duplication is also required. An extension of this task added all possible duplications of the correct responses. Other variants are "Kirkman's Problem" and bridge arrangement problems (**2**, p. 237; **31**, pp. 114–115; **2**, Ch. 10).

9. Using a series of city maps of uniform block design, a method of computing the maximum number of ways, as short as the shortest, of traversing these square and rectangular maps from the NW to the SE corners from knowledge of the number of blocks in rows and columns is required.

10. Using levers of the first class placed in equilibrium (**27**),

or the correlation array with the mean lines drawn in, the problem is developed for the re-discovery of the principle of moments.

11. Using a series of domino sets of the ordinary rectangle of two squares, or the equilateral triangle with three similar triangles, the generalization of the number of dominoes from knowledge of the number of dots permitted is required (**18**).

12. An extension of the problem of generalized dominoes, called "Designing Pastimes," involves introducing arbitrary restrictions of patterns or boundaries (**2**, pp. 64, and 255; **18**).

13. A factor stencil problem has been described in detail in Appendix A (**17**).

14. The re-discovery of Eratosthenes Sieve for sorting out prime numbers has also been assigned (**2**, **17**).

IV. *Tasks involving progression or power series:*

15. Using the Disc Transfer Puzzle (**2**, p. 288; **31**, pp. 31–32), a method of finding the fewest possible number of moves from knowledge of the number of discs in a pile is required. A supplementary principle describing a method of moving accurately is also required.

16. The Chinese Ring Puzzle supplies material for a similar set of generalizations, and possibly other puzzles of Ruger's list could be readily adapted for the construction of similar series (**2**, **33**).

V. *Other tasks essentially mathematical in type:*

17. A short cut method of squaring numbers ending in five has been described in the text of this article.

18. Other short-cut methods are usable, such as shifting from the decimal to the duo-decimal system of notation (**1**), or from one of several systems of measurement to another.

19. The water-dipping ingenuity test has received mention in the text (**36**, **37**). See also "Bachet's weight problems" (**2**, p. 34).

20. The checker puzzle problem (**9**, p. 34; **2**, pp. 77–80) provides a generalization for the number of moves and jumps from knowledge of the number of checkers used, and one for checking accuracy in moving and jumping.

21. Employing the familiar tabular device and utilizing an arithmetical series for the stubs of the rows, the relationship

between the rows and a column of answers, affords many possible generalizations, such as reciprocals, $N^2 - N$, etc.

22. The number series device in appropriate sets is usable.

VI. *Games involving uniformities and complete solutions:*

23. A series based upon John C. Peterson's study has been described in detail in Appendix A (32).

24. The Game of Nim (4) gives the initial impression of extreme difficulty. In this game, two or three piles of counters are placed upon a table. Two opponents play alternately and are permitted to remove one or more counters from one and only one pile at any play. The object of the game is to leave safe combinations for the opponent in order to win by removing the last counter. A rule is sought for determining how to compute the safe combinations.

VII. *Miscellaneous tasks:*

25. A bent wire puzzle series is given brief mention in the text (*supra*, p. 93).

26. Using artificial language constructs and giving vocabularies and translations, the learning of the grammatical and syntactical rules is required.

27. Using a series of figures for continuous tracing without raising the pencil, crossing, or re-tracing lines, a generalization may be required which will tell what kind of figures are unicursal, which are not, or where to begin and end such figures (2, Ch. 9; 9, p. 37). Variants include the puzzle of moving pawns along straight lines from some vacant corner to fill one less than the number of points; and the generalized order of travel that enables one to visit once and only once each point of a plane projection of each of the regular solids.

28. A series of chess or peg removal problems involving jumping and removing pieces may be constructed and the appropriate generalized procedures required (2, p. 109 ff.).

29. A series of block moving puzzles involving space restrictions and specific start and end patterns is similarly usable (2, p. 224).

30. A series of plane match puzzle patterns involving specific beginning and end patterns is likewise usable.

COMPARATIVE STUDIES OF FULL AND MIXED BLOOD NORTH DAKOTA INDIANS [1]

C. W. Telford
UNIVERSITY OF NORTH DAKOTA

The early comparative studies of Indian-white mixtures in America uniformly reported superior mental test performances of mixed as compared with full blood Indians (1, 2, 3, 4, 6). The tests used in these investigations were principally standard group intelligence tests of the language type, which reflect very markedly the different social, cultural, and educational backgrounds of the subjects. In a more recent study of North Dakota Indians (8) such differences between mixed and full bloods were not found. In this investigation the Peterson Rational Learning Test was used, an ideational learning test which seems to draw little on past experience and training, and which stimulates the subjects to approximately maximal effort throughout the performance. In addition to the learning test, certain performance tests of the Pintner-Paterson Series (Mare and Foal, Healy Puzzle "A," and Two-Figure Form Board), and one group performance test (The Goodenough "draw a man" Test) were administered. When comparisons were made between the test performances of the full and various mixed blood groups on the basis of these test performances, no significant differences appeared.

[1] The data from Wahpeton presented in this report were obtained by Miss Ingaborg Jonasson at the Wahpeton Indian School, Wahpeton, North Dakota, and will be found more in detail in an unpublished Master's Thesis on file in the University of North Dakota Library. Mr. Meinick Bodahl at the Fort Peck Indian School on the Fort Peck Indian Reservation in Eastern Montana obtained the data on the Fort Peck Indians. These are similarly included in an unpublished Master's Thesis in the University Library.

The individual test results from Fort Totten were obtained by the writer, while the group tests were given by WPA workers working under the supervision of Mr. C. H. Beitzel, principal of the school.

The Wahpeton and Fort Totten data were obtained as part of WPA Projects Nos. 987 and 1876, sponsored by the University of North Dakota and supervised by the writer. These data will also be included in a more extensive report of the entire projects which will be available at a later date.

This suggested the possibility that the alleged superiority of the mixed over full bloods was due to greater contact with white people and a white culture, with a resulting increased familiarity with the materials and information involved in the group intelligence tests. These differences might be accentuated by a higher social status and by the greater educational opportunities possibly enjoyed by the mixed bloods. Such differences should consequently not appear in the test performances of the two groups on performance tests and learning tests which are not dependent upon specific information previously acquired by the subjects. In the present study we have checked on this hypothesis, using a wide variety of tests.

Statement of Problem

More specifically our study involves a comparison of mixed and full blood Indians by means of tests of the following types: (1) Achievement tests of the educational type; (2) standard group intelligence tests of the linguistic type; (3) individual learning tests ('non-informational'), and (4) performance tests, both group and individual.

If the alleged superiority of the mixed bloods is due to their greater contact with whites and to the acquisition of information relative to such an environment, rather than to an inherent superiority of the white blood from which they have sprung, we would expect to find a decided superiority of the mixed bloods on the educational achievement tests and intelligence tests of an informational type (groups 1 and 2). This difference should decrease in extent or disappear entirely when learning and performance test results are used as a basis of comparison (groups 3 and 4). This is the hypothesis and general plan of the present study. In addition to this main problem, we have also obtained some data on the socio-economic levels of our Indian groups.

Method of Procedure

1. *Subjects.* The subjects of the present study were students of the various Indian schools of North Dakota and of one school on the Fort Peck Indian Reservation in Eastern Montana. North

Dakota schools include five on the Fort Berthold, three on the Turtle Mountain, and two on the Standing Rock reservations, as well as three boarding schools not on reservations. These three are located at Wahpeton, Fort Totten, and Bismarck, North Dakota. All of these are federal schools except the Cannon Ball School on the Standing Rock Reservation. This is a district school; but its enrollment is 98% Indian and consequently it has been included with the others as an "Indian" school. The pupils of these schools are principally Sioux and Chippewas, with a sprinkling from several other tribes. A previous study (8) has shown no differences in test performances between representatives of these two tribes. Hence the two groups are not treated separately.

As the data from all of the schools are not yet available, the present study will be based on the results obtained from three schools only. These are the Wahpeton (W), Fort Totten (F. T.) and Fort Peck (F. P.) schools respectively.

2. *Tests.* For our achievement tests we have used the New Stanford Achievement Tests, Form V, Primary and Advanced, at Wahpeton and Fort Totten, and the Torgeson Achievement series at Fort Peck.

The following Group Linguistic Intelligence Tests were used: The Otis Tests, Primary, Advanced, and S-A at Fort Totten and Wahpeton, and the Kuhlmann-Anderson Test at Fort Peck.

The individual ideational learning test was the Peterson Rational Learning Test. For a detailed description of this test see (7) and (8). It was not administered at Fort Peck; consequently our results are based on the Fort Totten and Wahpeton groups only.

For the individual performance tests we have used the Seguin and Casuist Form Boards and certain of the Pintner–Paterson series. These included the Seguin Form Board, Casuist Form Board, and Feature Profile for the Wahpeton children; and the Mare and Foal, Two Figure Form Board, and Healy Puzzle "A" for the Fort Totten and Fort Peck groups.

The group performance test was the Goodenough Intelligence Test which is based on the ratings of the drawing of a man.

All of these tests were administered and scored according to standard procedures. The prescribed time limits and instructions were strictly adhered to at all times. All papers were scored twice and all calculations checked for accuracy.

The degrees of Indian blood represented by the subjects were obtained from the government records. The reliability of these records is questionable but they are the best available. It is generally believed by those in charge of these records that those listed as full bloods are probably all Indian; the degree of mixture of the others is often uncertain. As a result our principal comparisons are between the full blood groups on the one hand and the entire mixed blood group on the other. In a few cases, where large numbers are represented, we are making finer distinctions within the latter group. It will be noted that as a statistical fact, any lack of reliability of the placements will tend to decrease the extent of the obtained differences. Consequently, any errors in placement increase rather than decrease the significance of those differences found to be statistically significant.

The group tests were administered to all students to whom they were applicable and who were in attendance on the days of the tests. We have followed our custom of giving the individual tests to all twelve-year-olds irrespective of grade placement.

Results

Table 1 summarizes the results obtained from the Stanford and Torgeson Achievement Tests. In this and all succeeding tables (W) refers to the Wahpeton group, (F. T.) to the Fort Totten group, and (F. P.) to the Fort Peck group. Positive values indicate that the mixed bloods excel while negative values show full blood superiority. Unless a minus sign appears before a figure, the value will be assumed to be positive.

The Fort Totten and Wahpeton scores are given in terms of "educational ages," for groups of equivalent chronological ages, while the Fort Peck data represent raw scores or "points." The Stanford figures are based upon total scores for the entire battery of tests in the series. Since no such total score is computed for

the Torgeson series, we have given median scores for each of the individual sub-tests.

This table shows that in every case the mixed bloods excel the full bloods in terms of median scores on the achievement tests. The critical ratios vary from .25 to 7.79. Averaging the ratios of the various sub-tests of the Torgeson series we obtain 2.71 as

TABLE 1

COMPARISONS OF ACHIEVEMENT TEST SCORES OF FULL AND MIXED BLOOD INDIANS

Test	Basis of Comparison	Medians: Mixed Bloods	Medians: Full Bloods	Diff. / P.E. of D.	Diff. / Q. of M.
Stanford Achievement Adv. (Fort Totten)	E.A.	11.69	10.58	4.27	1.17
Stanford Achievement Adv. (Wahpeton)	E.A.	13.42	12.18	4.59	1.14
Stanford Achievement Primary (Wahpeton)	E.A.	11.09	10.00	7.79	2.60
Torgeson Ach. Test (Fort Peck)	Points				
Reading		41.66	33.75	2.49	.90
Arithmetic Computation		45.00	23.00	4.78	1.79
Arithmetic Reasoning		22.50	15.00	2.30	.78
Language Usage		40.00	22.00	4.68	1.60
Spelling		62.50	61.00	.25	.10
Grammar		27.50	20.62	1.83	.85
History		31.25	16.25	2.94	1.84
Geography		33.75	22.50	2.35	1.33
Nature Study		41.25	30.00	2.05	1.16
Health		40.00	23.33	3.46	1.66

the mean of the group. Averaging this with the corresponding values for the Stanford Achievement Series, we obtain 4.84 as the mean critical ratio for all of our achievement tests. This ratio is well above the traditional standard of statistical reliability. Making the corresponding calculations for the last column, which gives the differences in terms of a standard unit (the quartile of the mixed group), we find that the mixed excel the full bloods by 1.53 quartile units, on the average. From this it seems quite clear that in school achievement the mixed bloods are definitely superior to their full blood classmates.

Table 2 shows the results obtained from the various group intelligence tests. All comparisons in this table are in terms of I. Q's. We recognize that I. Q's. are not direct indices of per-

formance and that the I. Q's. obtained from different tests are not always comparable, but since our primary interest here is not in inter-group comparisons but rather in intra-group comparisons, which are always in terms of the same test, it seemed that these values might be used. We have no way of determining whether the differences shown among our various groups taking different tests are due to differences in these tests, or to differences in the groups studied, or to both these factors. But this is not pertinent to our present problem.

TABLE 2

COMPARISONS OF INTELLIGENCE TEST SCORES OF FULL AND MIXED BLOOD INDIANS

Test	Median I.Q.'s Mixed Bloods	Median I.Q.'s Full Bloods	Diff. / P.E. of D.	Diff. / Q. of M.
Otis Primary (F. T.)	99.75	90.50	4.22	1.22
Otis Advanced (F. T.)	90.81	82.37	3.67	1.15
Otis S-A (Wahpeton)	89.31	81.25	3.95	.93
Otis Primary (Wahpeton)	94.79	89.15	2.56	.66
Kuhlmann-Anderson (Fort Peck)	82.50	80.57	.68	.29

An examination of Table 2 again shows a superiority of the mixed over the full blood groups in every case and with all tests. The mean critical ratio is 3.08 and the average difference in terms of the quartile of the mixed groups is .85. It will be noted that only one of these differences is statistically reliable, and that the average of the critical ratios is smaller than the corresponding values on the achievement tests (3.08 as compared to 4.82). The differences between the medians in terms of standard units are also much smaller (.85 as compared to 1.53 quartile units). Thus it seems that on the basis of our group linguistic intelligence tests the mixed bloods still excel the full bloods in test performance but not so markedly as with the achievement tests.

Table 3 presents comparable data from a single individual learning test, the Peterson Rational Learning Test. This test is scored in three ways: first, the total time required to do the learning; second, the number of erors made; and third, the number of repetitions required to complete the task. A different

story is told by this table. It will be noted that a high value represents a poor score (long time and more repetitions to learn, as well as more errors made). Consequently, the group with the lowest score numerically excels the other. In three of the six comparisons the mixed bloods excel (time and repetitions with the Fort Totten group and repetitions for the Wahpeton students) while the full bloods excel in three (errors for the Fort Totten and time and errors for the Wahpeton groups respectively). The average critical ratio is .15. The mean of the differences in terms

TABLE 3

COMPARATIVE SCORES OF MIXED AND FULL BLOOD INDIANS ON THE PETERSON RATIONAL LEARNING TEST

	Medians			
Basis of Comparison	Mixed Bloods	Full Bloods	Diff. / P.E. of D.	Diff. / Q. of M.
Time (Sec.) (Fort Totten)	204.13	237.50	.23	.24
Errors (Fort Totten)	22.50	21.66	—.06	—.05
Repetitions (Fort Totten)	7.20	12.00	2.08	1.68
Time (Sec.) (Wahpeton)	216.64	202.08	—.69	—.26
Errors (Wahpeton)	34.85	32.83	—.73	—.13
Repetitions (Wahpeton)	8.57	9.08	.05	.21

of the quartile of the mixed groups is .28. These values indicate that no significant difference exists between these groups on the basis of their performances on the Peterson Rational Learning Test.

Table 4 gives the results of the application of the various individual performance tests to all the twelve-year olds in these three schools. The picture shown here looks considerably like that of Table 3 except that the differences are shifting more and more to favor the full bloods. Out of fifteen comparisons, eleven favor the full blood groups, one shows no difference, and only three favor the mixed bloods. The mean critical ratio is —.82, and the average difference in quartile units is —.16. Although these differences are small and statistically unreliable, they appear in striking contrast to the results for the achievement and group intelligence tests shown in Tables 1 and 2. The decided difference in favor of the mixed bloods on the achievement tests was found to be greater than the corresponding difference for the

TABLE 4

COMPARISONS OF INDIVIDUAL PERFORMANCE TEST SCORES OF FULL AND MIXED BLOOD INDIANS

Test	Basis of Comparison	Medians: Mixed Bloods	Full Bloods	Diff. / P.E. of D.	Diff. / Q. of M.
Seguin Form Board	Time (W.)	19.85	18.16	—.10	—.29
	Errors (W.)	4.21	2.91	—4.06	—.84
Casuist Form Board	Time (W.)	144.71	127.08	—.70	—.24
	Errors (W.)	8.90	6.75	—1.36	—.46
Feature Profile	Time (W.)	116.64	102.08	—0.69	—.26
Mare and Foal	Time (F. T.)	26.25	31.66	1.11	.66
	(F. P.)	40.83	39.37	—.43	—.11
	Errors (F. T.)	.50	3.20	—4.90	—1.47
	(F. P.)	1.65	1.56	—.24	—.06
2-Figure Form Board	Time (F. T.)	36.66	50.00	1.56	.92
	(F. P.)	58.33	52.50	—.18	—.14
	Moves (F. T.)	11.16	9.20	—1.03	—.52
	(F. P.)	12.83	15.17	—1.31	—.46
Healy Puzzle "A"	Time (F. T.)	50.00	50.00	.00	.00
	Moves (F. T.)	10.40	10.66	.06	.02

group intelligence tests. The latter difference was similarly larger than the one based on the Peterson Rational Learning Test. On the individual performance tests the superiority of the mixed bloods has entirely disappeared and been replaced by a slight difference in favor of the full bloods.

Turning to the last table (Table 5) we find that this tendency is still further accentuated. This table makes comparisons on the basis of a single test, the Goodenough Test for measuring intelligence by means of the drawing of the figure of a man. The data of the Wahpeton group are recorded in terms of I.Q.'s; the other results are based upon raw scores. An examination of this

TABLE 5

COMPARISONS OF TEST SCORES OF MIXED AND FULL BLOOD INDIANS ON THE GOODENOUGH INTELLIGENCE TEST

Basis of Comparison	Medians: Mixed Bloods	Full Bloods	Diff. / P.E. of D.	Diff. / Q. of M.
I.Q. (Wahpeton)	93.55	99.00	—1.41	—.43
Points (Fort Totten 1937)	40.22	41.17	—.90	—.11
Points (Fort Totten 1932)	32.85	35.85	—2.61	—.30
Points (Fort Peck)	37.50	41.25	—1.60	—.33

table discloses the rather startling fact that, in every case, the full bloods excel the mixed bloods. This completely reverses the trend of the differences shown in Tables 1 and 2. The mean critical ratio in Table 5 is —1.63 and the average difference is —.29 quartile units. Thus we see that although the differences in favor of the full bloods are greater than in Table 4, they are still statistically unreliable.

From these comparisons it seems clear that the decided differences found in favor of the mixed bloods with the achievement and group "intelligence" tests disappear entirely when comparisons are made on the basis of individual and group "performance" tests.

Practically nothing is known of the relative social, economic, and educational backgrounds of the mixed and the full blood Indians. However, the Indian agent and the principal of the two-teacher school on the Fort Peck Indian Reservation independently rated the families of the children of the school on their degree of self-support. A seven-point rating scale was used, and where differences were found in the placements of the two raters the mean of the two was taken as the rating. In no case did the two persons differ by more than one point in their placements, and all "one point" differences were in the three middle positions. The ratings ranged from entirely self-supporting at the one extreme to entirely dependent upon governmental and other aids at the other. The agent did not know the purpose of the ratings or what comparisons were to be made, at the time he made his judgments, while the principal did have such knowledge. When the median scale values of the two groups were compared, a difference of 1.58 scale units was found in favor of the mixed bloods. The critical ratio of this difference is 4.48 and the difference in terms of the quartile of the mixed group is 1.80.

There is little question as to the comparative socio-economic levels of Indians and whites as groups, when judged by white standards. For example: Jamieson and Sandiford (5) found the Indians which they studied to show an average score of 13 against a norm of 56 on the Chapman Socio-economic Scale.

They did not compare mixed and full blood Indians. They also found that the Indian children attend school less regularly than the neighboring white pupils, the average attendance in percentage of total enrollment being 62.23 as contrasted with 67.14 for the whites. An age-grade table also shows them to be much more retarded than whites.

Garth (3), using only opinions based on the general impressions of persons having worked with the Indians for considerable periods of time, concludes that the general environmental, educational, and linguistic backgrounds of the homes favor the mixed as compared with the full blood Indians.

These ratings and opinions indicate that the mixed bloods are probably better off economically, educationally, and socially. We hope to be able to obtain more complete and objective data on the relative socio-economic and educational backgrounds of our subjects at a later date.

Limitations of space and incomplete analysis of data at the present time prevent the complete presentation of evidence, but the following additional conclusions seem warranted:

1. Full bloods are consistently more retarded in school than are mixed bloods. Mixed bloods of the same school grade are approximately one and one-half years younger. This is in keeping with the poorer showing on achievement tests when comparisons are made on the basis of chronological age.

2. The Indian groups tend to score below white norms for corresponding ages on the achievement and group intelligence tests.

3. The inferiority on the part of the Indians increases with increase in chronological age. In other words, with the Otis and Kuhlmann-Anderson Tests, the median I.Q.'s decrease with increase in age. Similarly with the Stanford and Torgeson Achievement Tests, E.Q.'s tend to decline with increasing chronological age. Of course, in the upper grades the possible graduation of the brighter pupils from school must be considered along with educational and general environmental factors as possible causes of these declines in performance with increase in age.

4. The Goodenough Drawing Test does not show any such

consistent change with chronological age. This suggests that educational and general environmental factors, rather than selection in the school system, may be the responsible factors in the declining I.Q.'s and E.Q.'s noted above.

5. On the Stanford Achievement and Otis Intelligence Tests comparisons of those $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$, and $\frac{4}{4}$ Indian show slight decreases in test scores with increases in degree of Indian blood. This further supports the data presented in Tables 1 and 2 where comparisons are made between the full and mixed bloods only.

6. Similar analysis of the results on the Goodenough Test for the Wahpeton group (the only group for which this analysis has been made to date) show the median I.Q.'s of the four groups to be as follows: 91.50 for $\frac{1}{4}$ bloods, 91.50 for $\frac{1}{2}$ bloods, 96.50 for the $\frac{3}{4}$ bloods and 99.00 for the $\frac{4}{4}$ bloods.

Summary and Conclusions

The hypothesis with which we began the present study is, briefly, as follows: if the apparent mental superiority of the mixed over full blood Indians, found by early workers on the problem, is due to their greater familiarity with the English language and with information based upon a white culture, we would expect to find them definitely superior on tests of an informational sort. We would similarly expect this difference to disappear when tests which are non-informational, non-linguistic, and largely "performance" in nature are used as the basis for comparisons.

The data presented in the present study definitely support this hypothesis. When comparisons between our mixed and full blood groups are made on the basis of test performances on the Stanford and Torgeson Achievement Tests, the mixed bloods are decidedly and consistently superior. On the group linguistic intelligence tests (Otis and Kuhlmann–Anderson) they are still superior by statistically reliable amounts, but the differences are smaller than with the achievement tests. With the non-informational, ideational learning test (Peterson Rational Learning Test) the two groups are practically equal. On the individual performance tests (Seguin Form Board, Casuist Form Board,

and four tests of the Pintner–Paterson series) the differences obtained are in general, statistically unreliable, but tend to favor the full bloods. With the group performance test (Goodenough Draw-a-Man Test) the full bloods are still more superior, although the difference again is not reliable.

In other words, as the tests become more and more of the informational and achievement nature, the differences increasingly favor the mixed blood; or, conversely, as the tests depend more and more on basic learning and manipulative abilities, the differences between the two groups tend to disappear. There is, correspondingly, some evidence that the mixed bloods are better off economically and socially than are the full bloods. These facts indicate that the apparent superiority of the mixed over the full blood Indian is environmentally rather than biologically determined.

In general the results of the present study are in keeping with those of our earlier report to which reference has already been made (8). Both of them have demonstrated no significant variations in test scores with changes in degree of Indian blood, for the learning and performance tests. They also agree with those studies which have found such correlations when ordinary standard intelligence tests are used. Several of the latter studies have been published.

The first seems to have been that of Hunter and Sommermier (4) who gave the Otis Intelligence Test to 715 American Indians in attendance at the Haskell Indian Institute at Lawrence, Kansas. These students came from 65 tribes and represent all degrees of mixtures. When the subjects were classified into four groups (1/4, 1/2, 3/4, and 4/4) according to the degree of Indian blood, a steady decrease in average scores was found with increase in Indian blood. The median I.Q.'s of these respective groups were 109, 91, 78 and 67. These differences were found to be statistically reliable. It will be noted that most of these differences are rather large, much larger in fact, than those found by any later studies, including the present one. Whether the time factor (1922 to 1937) has anything to do with it, we do not know. It is possible that the differences between the groups

in the extent of their contacts with whites are becoming less as time goes on. This would tend to reduce the differences, which, as we have shown in the present study, seem to be largely the result of such contacts.

In the same year (1922) Garth (2) made comparisons of mixed and full blood Indians of southwestern United States. His comparisons were between groups in the same school grades by means of a series of association and word-building tests. All of the tests were linguistic in nature. He found that the mixed blood group is eleven per cent better than the full blood group in performance, in spite of the fact that the latter are older chronologically than the former. We have found similar results for our language tests.

Fitzgerald and Ludeman (1) gave the National Intelligence Test to 41 Indian children in St. Mary's Mission School, Springfield, South Dakota, and the Otis Group Intelligence Scale to 42 high school students in Santee Normal Training School at Santee, Nebraska. They divided the subjects into five groups on the basis of degree of Indian blood and found the median I.Q. for each to be: $1/5$ or less Indian blood, 93; $2/5$, 94; $3/5$, 88; $4/5$, 89; $5/5$, 89. These differences are more nearly the size of those found in the present study, but, in view of the small numbers of subjects represented, are of doubtful reliability.

Paschal and Sullivan (6) report "that there is a direct (negative) relationship between the amount of Indian blood and the social, school, and mental status of Mexicans in Tucson." Much of the material presented by these authors is not quantitative and it is consequently difficult to determine the extent and reliability of this relationship from their report.

Jamieson and Sandiford (5) in a study of Southern Ontario Indians report findings somewhat comparable to our own. They administered a wide variety of tests, some of which were used in the present study, to a large number of these Indians. The primary interest of these investigators was in the intellectual status of the Indians as compared with other races. Consequently, they present very little data on the comparative performances of mixed and full bloods. They do make the

statement, however, that " the I.Q. seems to rise with the admixture of white blood, but the results must be interpreted with caution owing to the fact that the numbers of half-bloods and three-quarter bloods (apparently the only groups compared) are few and also because the amount of white blood cannot be determined with accuracy " (5, p. 319). Their data show the above statement to be true of some tests, notably the National Intelligence Test, but not of otheres i. e., the Pintner–Paterson Performance series. These results agree with our own.

So far we have found no studies which conflict with the results and conclusions of the present investigation. We shall soon have available the results of the entire investigation, which includes more tests and many more subjects than those reported here. We do not anticipate, however, that the more complete data will markedly alter the trend of results reported in the present study.

Bibliography

1. Fitzgerald, J. A., and Ludeman, W. W. The intelligence of Indian children. *J. Comp. Psychol.*, 1926, **6**, 319–328.
2. Garth, Thomas R. A comparison of mental abilities of mixed and full blood Indians on a basis of education. *Psychol. Rev.*, 1922, **19**, 221–236.
3. Garth, Thomas R. A comparison of the intelligence of Mexican and mixed and full blood Indian children. *Psychol. Rev.*, 1923, **30**, 388–401.
4. Hunter, W. S., and Sommermier, E. The relation of degree of Indian blood to score on the Otis Intelligence Test. *J. Comp. Psychol.*, 1922, **2**, 257–275.
5. Jamieson, E., and Sandiford, P. The mental capacity of Southern Ontario Indians. *J. Educ. Psychol.*, 1928, **19**, 313–328 and 536–551.
6. Paschal, F. C., and Sullivan, L. R. Racial differences in the mental and physical development of Mexican children. *Comp. Psychol. Monog.*, 1925, **3**, No. 14. Pp. 76.
7. Peterson, J., and Telford, C. W. Results of group and of individual tests applied to the practically pure-blood children of St. Helena Island. *J. Comp. Psychol.*, 1930, **11**, 115–144.
8. Telford, C. W. Test performance of full and mixed blood North Dakota Indians. *J. Comp. Psychol.*, 1932, **14**, 123–145.

COMPARATIVE STUDIES OF CERTAIN MENTAL DISORDERS AMONG WHITES AND NEGROES IN GEORGIA DURING THE DECADE 1923–32 [1]

J. E. Greene

THE UNIVERSITY OF GEORGIA

This paper has two major purposes. First, it proposes to make comparisons, disregarding race, for several disorders and disorder groupings with respect to the following factors: (a) age of patients at the time of first admission; (b) distribution of first admissions within given clinical categories expressed as percentages of total first admissions, by age levels; (c) standard rates of first admission, disregarding age levels; (d) deaths per 100 resident patients; (e) duration of hospital residence prior to death for patients dying at the hospital; (f) age at the time of death of patients dying at the hospital. The second purpose is that of making pertinent interracial comparisons with respect to each of the above factors, in each of the several clinical categories selected for special study. Due to space limitations it has not been possible to make complete analyses of any of these factors. In view of the fact that first admissions perhaps constitute the

[1] The author is indebted to many persons who coöperated to make possible the studies reported here. Mrs. Mary Bondurant and Mrs. Dorothy Whitehead, National Youth Administration Supervisors of The University of Georgia, have each generously aided by assigning typists, statistical clerks, and calculating machine operators to work under the author's supervision in tabulating certain source data and at other routine tasks. Certain of the more intricate calculations herein reported have been double-checked by the following graduate students: Mr. J. L. DuPree, Mr. J. S. Jacob, Miss Martha Moore, Mr. R. T. Osborne, and Miss Hazel Usry. Valuable secretarial aid in the preparation of the manuscript has been given by Miss Kay Johnson and Miss Hazel Usry. Most of the primary source data were made available through the courtesy and coöperation of Superintendent J. W. Oden of the Milledgeville State Hospital, Milledgeville, Georgia. Each of these individuals is due much appreciation for his aid, without which it would have been difficult or impossible to carry out this study.

most adequate single criterion of the incidence of mental disorders, it has been considered advisable to attempt fairly detailed analyses of the ages of first admissions, by psychosis, by race.

In securing the data bearing on these factors the general procedure was as follows: Selected data from the Annual Reports of the Georgia State Hospital for the Insane were tabulated for each of the years 1923–32, inclusive, and averages were computed for the decade. Since it was clearly impractical to make separate analyses of every clinical category in use during the decade under study, it was decided to select six representative psychoses for special study, as follows: (1) senile psychoses, (2) psychoses with cerebral arteriosclerosis, (3) general paralysis, (4) psychoses with cerebral syphilis, (5) manic-depressive psychoses, and (6) dementia praecox psychoses. It is to be observed that the above six psychoses may be combined into sub-groups of disorders, to wit: (1) and (2) refer to disorders commonly associated with advanced chronological age; (3) and (4) refer to disorders accompanied by syphilitic infection; (5) and (6) refer to disorders which are generally regarded as being of a "functional" nature (i.e., presumably unaccompanied by pathognomic "organic" symptoms). Since each of the first four of the above six disorders is considered to be of an "organic" nature, it is possible to combine them into a sub-group of "organic disorders." In addition to the six specific psychoses and the several sub-groups referred to above, a further category of data has been utilized, namely, "all clinical groups" (*i.e.,* the sum total of all disorders of every kind occurring among patients in the Georgia State Hospital for the Insane).

The above exposition should serve to explain the following twelve-fold classification of data utilized in the twelve columns of all tables in the study: Column 1, senile psychoses; Column 2, psychoses with cerebral arteriosclerosis; Column 3, general paralysis; Column 4, psychoses with cerebral syphilis; Column 5, manic-depressive psychoses; Column 6, dementia praecox psychoses; Column 7, both old-age disorders (senile psychoses and psychoses with cerebral arteriosclerosis); Column 8, both syphilitic disorders (general paralysis and psychoses with cerebral

syphilis); Column 9, all four organic disorders (senile psychoses, psychoses with cerebral arteriosclerosis, general paralysis, and psychoses with cerebral syphilis); Column 10, both functional disorders (manic-depressive psychoses and dementia praecox psychoses); Column 11, all six disorders (the sum of Columns 1, 2, 3, 4, 5, and 6); Column 12, all clinical groups (the sum of all patients of every type of disorder in the Georgia State Hospital for the Insane).

A brief statement relative to the nature and reliability of the present sampling would seem to be appropriate. Only one hospital for insane patients is maintained by the state of Georgia. There are several small, privately-owned hospitals for mental patients within the state, but the total number of resident patients in these private hospitals is small (approximately 2 per cent of the state total). The State Prison Commission reports that no mental patients are kept in the state prisons or in the county jails. There are no county or municipal hospitals for mental patients. It is obvious, therefore, that the present data include practically all hospitalized mental patients in Georgia during the decade under study. No reliable data are available relative to equality of opportunities for hospitalization for the two races or for different types of mental patients. Georgia law requires that all patients must be adjudged insane by a Lunacy Commission in the county of legal residence, before being legally committed to the State Hospital by the County Ordinary. It has not been possible to determine whether this legal requirement tends to limit the number of commitments among certain types of mental disorders or to make for racial differences in commitment rates.

A definition of terms used in referring to certain classes of patients is perhaps in order. By *first admissions* is meant all patients entering the hospital for the first time. The term *resident patients* refers to hospital inmates, whether first admissions or readmissions, under treatment as of a given date. As herein used, the term *deaths* refers to patients dying during a period of hospital residence.

Little comment seems necessary concerning the statistical procedures employed herein. With the exception of Table 3 (in

which the arithmetic mean is employed), the median is used as the measure of central tendency. Quartile deviations of distribution are used as measures of absolute dispersion. In using coefficients of variation as measures of relative variability the following formula was employed: $V = \frac{100(Q_3 - Q_1)}{(Q_3 + Q_1)}$. In the analyses of first admissions it was found possible to make direct comparisons of the two races and among each of the several clinical categories themselves on the basis of the average annual number of first admissions *per 100,000 general population* of the same race. The weighted numbers thus obtained are referred to as standard rates of first admission. The figures on the general population used in computing the standard rates for each race, respectively, and for both races combined, were obtained from the 14th (1920) and 15th (1930) Census Reports of the United States. At the base of Table 3, a direct comparison is made of whites and negroes in terms of negro-white ratios. These ratios may be thought of as percentages (i.e., per cent negroes are of whites) or as the number of negroes per standard number (the numbers 100, 10 or 1 may be used as the standard number) of whites. Similar ratios are utilized in comparing the several clinical categories among themselves with respect to any given classification of data.

The arrangement of the tabular material is designed to facilitate the two types of comparisons previously mentioned: (a) comparison of the several clinical categories among themselves; (b) comparison of whites and negroes within each clinical category. The evaluation of our data relating to the several general factors mentioned in the opening paragraph of this paper can perhaps best be achieved by means of sub-analyses and topical discussions. Since more data are presented on the factor of age at first admission than on any of the other factors, the breakdown of the data on this factor will be correspondingly complex. In our discussion of this factor, analyses will be made of the following topics: summary comparisons of the ages of first admissions, and distribution of the first admissions among the several clinical categories according to age levels. Following the discussion of these

topics brief summary comparisons of the several clinical groups among themselves and of racial differences within such groups will be made with respect to the following: standard rates of first admission; comparative distributions of first admissions and of deaths; deaths per 100 resident patients; duration of hospital residence of patients dying during the decade; age of death of patients during the decade. In all analyses involving comparisons of the clinical categories among themselves it has been the practice to use the data on both races (listed under "TOTAL" in the various tables) as the basis of such comparisons. Interracial comparisons have involved the use of negro-white ratios as well as other indirect comparisons.

Summary Comparisons of the Ages of First Admissions

Summary comparisons of the ages of the first admissions in the several disorder groups are shown in Table 1.

Comparison of disorders, disregarding race.—The median age at the time of first admission for all clinical groups is 36.19 years. Of the six separate clinical groups studied, only dementia praecox and manic-depressive psychoses have lower median ages at first admission than that of all clinical groups, their median ages being 31.14 years and 34.08 years, respectively. Of the several disorder groups, senile psychoses have much the highest median age at first admission—68.91 years. The median ages at the time of first admission for each of the remaining groups are, in order of size, as follows: with cerebral arteriosclerosis, 62.58 years; general paralysis, 42.08 years; with cerebral syphilis, 41.33 years. It is interesting to note that among the four organic disorders, patients with syphilitic disorders (Table 1, Column 8) have a much lower median age at the time of first admission than do patients with "old age" disorders (Table 1, Column 7)—the median ages being 42.04 years and 65.03 years, respectively. If patients in the four organic disorders (Table 1, Column 9) are compared with those in the two functional disorders (Table 1, Column 10), it is found that the former are committed at 24.37 years older age than the latter—the median ages being 57.11 years and 32.74 years, respectively. The median

TABLE 1

MEDIAN AGES AT FIRST ADMISSION OF PATIENTS IN EACH OF THE TWELVE CLINICAL GROUPS, BY RACE, FOR GEORGIA, 1923–32

Quartile Deviations of Distributions, Probable Errors of Medians, and Coefficients of Variation are also indicated. At the base comparisons are made of Whites and Negroes on the basis of median ages at first admission. Critical ratios and chances in 100 of a true difference in medians are shown. Negative differences indicate that Negroes are older at the time of first admission. The numerals 1–12 under "Disorder Groups" refer to the following descriptive designations: (1) Senile; (2) Cerebral Arteriosclerosis; (3) General Paralysis; (4) Cerebral Syphilis; (5) Manic-Depressive; (6) Dementia Praecox; (7) Sum of Columns 1–2; (8) Sum of Columns 3–4; (9) Sum of Columns 1–4; (10) Sum of Columns 5–6; (11) Sum of Columns 1–6; (12) All Clinical Groups.

	Disorder Group (See legend)											
	1	2	3	4	5	6	7	8	9	10	11	12
Total												
Median	68.91	62.58	42.08	41.33	34.08	31.14	65.03	42.04	57.11	32.74	38.89	36.19
P.E. (dist.)	5.55	8.65	7.87	4.67	10.23	7.92	8.02	7.88	12.29	9.34	12.73	11.28
P.E. (median)	.39	.39	.38	.84	.27	.25	.30	.37	.36	.19	.21	.16
V	8.05	13.80	18.72	24.37	29.21	24.49	12.42	19.03	21.81	27.53	31.57	31.55
Whites												
Median	69.67	65.33	42.35	37.00	37.07	32.19	67.07	42.05	62.82	34.99	41.14	37.69
P.E. (dist.)	5.17	7.69	7.45	2.56	10.27	8.48	6.48	7.47	10.43	9.79	13.60	12.25
P.E. (median)	.44	.45	.60	.83	.35	.36	.31	.59	.43	.26	.30	.21
V	7.41	11.83	17.66	8.01	29.22	25.46	9.66	17.83	17.13	27.45	31.89	31.60
Negroes												
Median	66.62	57.76	41.90	43.50	29.58	29.29	59.82	42.02	51.22	29.76	36.36	33.99
P.E. (dist.)	6.72	8.74	8.21	11.50	8.72	7.46	8.91	8.37	11.13	8.09	11.57	11.20
P.E. (median)	.86	.61	.50	2.51	.37	.34	.54	.49	.47	.25	.29	.23
V	10.08	14.72	19.47	27.38	27.97	23.73	14.61	19.83	21.63	25.83	30.74	31.37
	Negro-White Comparisons											
Differences in medians	3.05	7.57	.45	—6.50	7.49	2.90	7.25	.03	11.60	5.23	4.78	3.70
Critical ratios	3.14	1.00	.57	2.46	14.69	5.80	11.69	.04	18.13	14.53	11.38	11.94
Chances in 100 of a true diff.	98	75	64	95	100	100	100	51	100	100	100	100

age at first admission for all clinical groups (36.19 years) is 2.70 years less than that for the six clinical groups combined (38.89 years).

Stated in order of size, the quartile deviations for each of the several clinical groups are as follows: all six disorders, 12.73 years; all four organic disorders, 12.29 years; all clinical groups, 11.28 years; manic-depressive, 10.23 years; both functional disorders, 9.34 years; with cerebral arteriosclerosis, 8.65 years; both old age disorders, 8.02 years; dementia praecox, 7.92 years; general paralysis, 7.87 years; both syphilitic disorders, 7.88 years; senile, 5.55 years; with cerebral syphilis, 4.67 years. The absolute variability in age at first admission is thus seen to vary considerably from one disorder group to another. Fifty per cent of the patients with psychoses with cerebral syphilis are first hospitalized betwen the ages of 36.66 years and 46.00 years, whereas the corresponding interquartile age range for manic-depressive first admissions is from 23.85 years to 44.31 years. The coefficients of variation likewise show marked differences among the several clinical groups in the relative variability in age at first admission. For example, the variability of all clinical groups (31.55) is roughly four times as great as that of senile psychoses (8.05).

Inter-racial comparisons.—Of twelve comparisons made in Table 1, with the single exception of psychoses with cerebral syphilis, negroes have lower median ages at the time of first admission than do whites; however, in this disorder negroes are 6.50 years older than whites at the time of first admission, and this difference is almost completely reliable. In seven of the eleven remaining comparisons, negroes are found to be appreciably and reliably younger than whites at the time of first admission. Negroes are slightly or appreciably younger than whites in the four remaining comparisons, but the differences lack complete statistical reliability.

The amounts of difference between the two races in median age at the time of first admission among the several clinical categories are, stated in order of size, as follows: all four organic disorders, 11.60 years; with cerebral arteriosclerosis, 7.57 years;

manic-depressive, 7.49 years; both old-age disorders, 7.25 years; with cerebral syphilis, —6.50 years (negroes are older); both functional disorders, 5.23 years; all six disorders, 4.78 years; all clinical groups, 3.70 years; senile, 3.05 years; dementia praecox, 2.90 years; general paralysis, .45 years; both syphilitic disorders, .03 years.

Comparisons of the relative variability of the two races in age at the time of first admission may be made in terms of the following negro-white ratios of coefficients of variation: senile, 1.36 to 1; with cerebral arteriosclerosis, 1.24 to 1; general paralysis, 1.10 to 1; with cerebral syphilis, 3.42 to 1; manic-depressive, .96 to 1; dementia praecox, .93 to 1; both old-age disorders, 1.51 to 1; both syphilitic disorders, 1.11 to 1; all four organic disorders, 1.26 to 1; both functional disorders, .94 to 1; all six disorders, .96 to 1; all clinical groups, .99 to 1. The amount of racial difference in relative variability in age at first admission is inconstant from one clinical category to another. Negroes are roughly three and one-half times as variable as whites in the case of psychoses with cerebral syphilis; while in the dementia praecox category, negroes are only 93 per cent as variable as whites.

Distribution of First Admissions Among the Several Disorder Groups According to Age Levels

Table 2 presents data showing characteristic groupings of the several disorders with respect to age levels. Summary analyses of the data in Table 2 indicate, in general, the following: In the youngest age intervals, manic-depressive psychoses constitute for both races a higher percentage of all patients than does any other single disorder. This finding is contrary to many others reported in the literature, in most of which dementia praecox is reported to constitute the highest percentage of first admissions at the earlier age levels. The middle age intervals are characterized by relatively high rates of incidence among general paralytics, manic-depressives and schizophrenics. In the oldest age intervals, psychoses with cerebral arteriosclerosis and senile psychoses predominate.

TABLE 2

Distribution of First Admissions Within Given Clinical Categories According to Per Cent of Total First Admissions, by Age Groups, by Race

The numerals 1–12 under "Disorder Groups" refer to the following descriptive designations: (1) Senile; (2) Cerebral Arteriosclerosis; (3) General Paralysis; (4) Cerebral Syphilis; (5) Manic-Depressive; (6) Dementia Praecox; (7) Sum of Columns 1–2; (8) Sum of Columns 3–4; (9) Sum of Columns 1–4; (10) Sum of Columns 5–6; (11) Sum of Columns 1–6; (12) All Clinical Groups.

	Disorder Group (See legend)											
Ages	1	2	3	4	5	6	7	8	9	10	11	12
Total												
Under 20			1.3	.2	26.5	13.7		1.5	1.5	40.1	41.6	100.0
20–29		.1	2.8	.4	28.8	28.2	.1	3.3	3.3	57.0	60.3	100.0
30–39		1.1	11.0	.5	28.1	23.7	1.1	11.5	12.6	51.7	64.3	100.0
40–49	1.8	14.7	13.1	.9	28.9	18.6	5.5	14.1	19.6	47.5	67.1	100.0
50–59	2.8	23.0	11.0	.5	24.0	9.1	25.8	11.5	37.3	32.9	70.2	100.0
Above 60	27.5	44.1	5.0	.6	9.3	2.5	71.6	5.6	77.2	11.8	88.9	100.0
Unknown	11.1	9.3	3.7		27.8	14.8	20.4	3.7	24.1	42.6	66.7	100.0
Total	3.7	9.1	7.5	.6	25.5	18.6	12.8	8.1	20.8	44.1	64.9	100.0
Whites												
Under 20			.6		21.3	13.3		.6	.6	34.6	35.1	100.0
20–29			1.7	.3	27.8	26.3		2.0	2.0	54.1	56.1	100.0
30–39		.5	7.0	.5	29.8	22.1	.5	7.4	8.0	51.9	59.9	100.0
40–49		2.5	8.4	.5	33.3	17.6	2.5	9.0	11.5	50.9	62.4	100.0
50–59	1.1	18.9	6.6		27.9	10.3	20.0	6.6	26.6	38.1	64.7	100.0
Above 60	29.4	43.9	2.2	.1	11.0	1.9	73.3	2.3	75.7	12.9	88.6	100.0
Unknown	21.1	10.5			15.8	21.3	31.6		31.6	36.8	68.4	100.0
Total	4.4	9.1	4.7	.3	26.2	17.2	13.4	5.0	18.4	43.4	61.8	100.0
Negroes												
Under 20			2.0	.5	32.1	14.1		2.5	2.5	46.2	48.7	100.0
20–29		.1	4.2	.6	30.0	30.4	.1	4.8	4.9	60.4	65.2	100.0
30–39		1.7	16.4	.6	25.8	25.7	1.7	17.0	18.8	51.5	70.3	100.0
40–49	.5	9.9	20.7	1.6	21.7	20.2	10.4	22.4	32.7	41.9	74.7	100.0
50–59	5.3	29.5	18.1	1.3	17.6	7.1	19.3	13.2	54.2	24.7	78.9	100.0
Above 60	22.8	44.7	11.6	1.6	5.1	3.9	67.5	13.2	80.7	9.0	89.7	100.0
Unknown	5.7	8.6	5.7		34.3	11.4	14.3	5.7	20.0	45.7	65.7	100.0
Total	2.7	9.2	11.6	.9	24.5	20.7	11.8	12.5	24.3	45.2	69.5	100.0

The distribution of disorders according to age levels may be considered in greater detail by dividing the life span into broad age intervals of ten or more years, as follows: under 20; 20–29; 30–39; 40–49; 50–59; above 60. It is evident that as one passes from younger to older age groups, the psychoses vary in their relative frequencies in characteristic fashion. The leading classification among patients under 20 years of age is manic-depressive, constituting 26.5 per cent of all patients of both races at this age level. The next highest category is dementia praecox, including 13.7 per cent of the total. There are no senile psychoses or psychoses with cerebral arteriosclerosis at this age level, and the number of first admissions with general paralysis (1.3 per cent) and psychoses with cerebral syphilis (.2 per cent) is negligible. It is to be noted that negroes have a consistently higher percentage of first admissions than whites at this age level, particularly among manic-depressives.

In the age level 20 to 29 years, the distribution of the six disorders, expressed as a per cent of the total for both races, is as follows: manic-depressive, 28.8 per cent; dementia praecox, 28.2 per cent; general paralysis, 2.8 per cent; psychoses with cerebral syphilis, .4 per cent; psychoses with cerebral arteriosclerosis, .1 per cent; senile, no first admissions. As was the case for the age level under 20 years, negroes have in each of twelve comparisons a consistently higher rate of first admissions at the age level 20 to 29 years.

In the age group 30 to 39 years, manic-depressive and dementia praecox psychoses are again the predominant disorders, including 28.1 per cent and 23.7 per cent, respectively, of the total first admissions of both races. General paralysis begins to be prominent, accounting for 11.0 per cent of the total. There are still no first admissions among seniles, and the rates for cerebral arteriosclerotics and cerebral syphilitics remain small. With the exception of manic-depressive psychoses, negroes continue to have higher rates of first admissions in the disorders studied than do whites. Stated differently, there is a pronouncedly greater tendency among negroes than among whites for the disorders

under study to constitute a larger proportion of all first admissions at this age level.

In the next age group, 40 to 49 years, the distribution of the six disorders, expressed as a per cent of total admissions for both races, is as follows: manic-depressive, 28.9 per cent; dementia praecox, 18.6 per cent; psychoses with cerebral arteriosclerosis, 14.7 per cent; general paralysis, 13.1 per cent; senile, 1.8 per cent; psychoses with cerebral syphilis, .9 per cent. Manic-depressive and dementia praecox psychoses continue to be predominant, but there is a sharp increase in the percentage of psychoses with cerebral arteriosclerosis as well as slight increases in each of the three remaining disorders. Only in the case of dementia praecox is there a decrease over the preceding age level. With the exception of manic-depressive psychoses, negroes continue to have higher rates of first admissions in the disorders studied than do whites.

The distribution of the disorders for both races combined in the age interval 50 to 59 years, stated in order of size, is as follows: manic-depressive, 24.0 per cent; with cerebral arteriosclerosis, 23.0 per cent; general paralysis, 11.0 per cent; dementia praecox, 9.1 per cent; senile, 2.8 per cent; with cerebral syphilis, .5 per cent. As contrasted with the preceding age interval, a marked increase is found in psychoses with cerebral arteriosclerosis and a marked decrease in dementia praecox. The remaining disorders show only slight changes from the distribution found at the age interval 40 to 49 years. Racial differences in distribution are noteworthy in several instances: psychoses with cerebral arteriosclerosis (whites, 18.9 per cent—negroes, 29.5 per cent); general paralysis (whites, 6.6 per cent—negroes, 18.1 per cent); manic-depressive (whites, 27.9 per cent—negroes, 17.6 per cent).

Further characteristic changes appear in the distribution of the disorders at the age interval above 60 years. Marked increases are found in senile psychoses and psychoses with cerebral arteriosclerosis, the rates being 27.5 per cent and 44.1 per cent, respectively. Significant decreases are noted in general paralysis, manic-depressive psychoses, and dementia praecox psychoses.

Racial differences are pronounced in the following categories: senile (whites, 29.4 per cent—negroes, 22.8 per cent); general paralysis (whites, 2.2 per cent—negroes, 11.6 per cent); manic-depressive (whites, 11.0 per cent—negroes, 5.1 per cent).

Various Supplementary Comparisons

Comparisons will next be made with respect to the following factors: standard rates of first admission; distributions of first admissions and of deaths; deaths per 100 resident patients. Table 3 shows standard rates of first admission, by clinical groups, by race; the average annual number of first admissions and of deaths, respectively, in each of the twelve clinical groups, by race; the distribution of first admissions and of deaths within each clinical category expressed as a per cent of first admissions and deaths, respectively, in all clinical groups; deaths per 100 resident patients, by clinical groups, by race. At the base of Table 3, negro-white ratios (obtained by dividing a given negro rate by the corresponding white rate) are shown for the following factors: standard rates of first admission (i.e., first admissions per 100,000 general population for each race separately); average annual first admissions (i.e., the gross number of first admissions, disregarding the proportion of the two races in the general population); average annual deaths (i.e., the gross number of deaths, disregarding the proportion of the two races in the hospital population); deaths per 100 resident patients for each race separately (i.e., standard death rates).

The standard rates of first admission for both races combined show marked variations from one disorder to another, as follows: senile, 1.11; with cerebral arteriosclerosis, 2.77; general paralysis, 2.28; with cerebral syphilis, .17; manic-depressive, 7.74; dementia praecox, 5.64. The standard rate for manic-depressives is thus seen to be more than forty-five times as high as that for cerebral syphilitics.

The negro-white ratios of standard rates of first admission for each of the several clinical categories are as follows: senile, .66 to 1; with cerebral arteriosclerosis, 1.09 to 1; general paralysis, 2.69 to 1; with cerebral syphilis, 3.22 to 1; manic-depressive,

TABLE 3

INTERCOMPARISONS AMONG EACH OF SEVERAL FACTORS FOR EACH OF THE TWELVE CLINICAL GROUPS, BY RACE, FOR GEORGIA, 1923–32

At the base is a direct comparison of Whites and Negroes on certain factors, expressed as a percentage (*i.e.*, per cent Negroes are of Whites on each of certain factors and for each clinical group) or as a Negro-White ratio (*i.e.*, number of Negroes per 100 Whites). The numerals 1–12 under "Disorder Groups" refer to the following descriptive designations: (1) Senile; (2) Cerebral Arteriosclerosis; (3) General Paralysis; (4) Cerebral Syphilis; (5) Manic-Depressive; (6) Dementia Praecox; (7) Sum of Columns 1–2; (8) Sum of Columns 3–4; (9) Sum of Columns 1–4; (10) Sum of Columns 5–6; (11) Sum of Columns 1–6; (12) All Clinical Groups.

	Disorder Groups (See legend)											
	1	2	3	4	5	6	7	8	9	10	11	12
Total												
Standard rates of 1st adm.	1.11	2.77	2.28	.17	7.74	5.64	3.88	2.44	6.32	13.38	19.70	30.34
Average annual first admissions	32.2	80.3	66.1	4.8	224.7	163.7	112.5	70.9	183.4	388.4	571.8	880.4
Average annual deaths	35.7	57.3	53.5	3.1	53.5	53.8	93.0	56.6	149.6	107.3	256.9	392.4
Distribution of 1st adm. by per cent	3.7	9.1	7.5	.6	25.2	18.6	12.8	8.1	20.8	44.1	64.9	100.0
Distribution of deaths by per cent	9.1	14.6	13.6	.8	13.6	13.7	23.7	14.4	38.1	27.4	65.5	100.0
Deaths per 100 resident patients	34.8	36.9	38.5	24.2	4.7	2.8	36.1	38.1	36.8	3.5	7.4	7.4
Whites												
Standard rates of 1st adm.	1.28	2.67	1.37	.09	7.70	5.03	3.94	1.45	5.39	12.73	18.12	29.33
Average annual first admissions	22.5	47.0	24.1	1.5	135.7	88.7	69.5	25.6	95.1	224.4	319.5	517.2
Average annual deaths	26.6	32.3	17.9	.9	26.9	27.7	58.9	18.8	77.7	54.6	132.3	207.0
Distribution of 1st adm. by per cent	4.4	9.1	4.7	.3	26.2	17.2	13.4	5.0	18.4	43.4	61.8	100.0
Distribution of deaths by per cent	12.9	15.6	8.7	.4	13.0	13.4	28.5	9.1	37.5	26.4	63.9	100.0
Deaths per 100 resident patients	40.1	34.4	22.6	39.1	3.9	2.2	36.7	23.0	32.1	2.8	6.1	5.9
Negroes												
Standard rates of 1st adm.	.85	2.92	3.69	.29	7.82	6.59	3.78	3.98	7.75	14.40	22.16	31.90
Average annual first admissions	9.7	33.3	42.0	3.3	89.0	75.0	43.0	45.3	88.3	164.0	252.3	363.2
Average annual deaths	9.1	25.0	35.6	2.2	26.6	26.1	34.1	37.8	71.9	52.7	124.6	185.4
Distribution of 1st adm. by per cent	2.7	9.2	11.6	.9	24.5	20.7	11.8	12.5	24.3	45.2	69.5	100.0
Distribution of deaths by per cent	4.9	13.5	19.2	1.2	14.3	14.1	18.4	20.4	38.8	28.4	67.2	100.0
Deaths per 100 resident patients	25.1	40.8	63.2	21.0	6.0	3.8	34.9	56.6	43.7	4.6	9.6	10.4
Negroes/Whites												
Standard rates of 1st adm.	66	109	269	322	102	131	96	275	144	113	122	109
Average annual first admissions	43	71	174	220	66	85	62	177	93	73	79	70
Average annual deaths	34	77	199	244	99	94	58	201	93	97	94	90
Deaths per 100 resident patients	62	119	280	54	154	168	95	246	136	164	157	176

1.02 to 1; dementia praecox, 1.31 to 1; both old-age disorders, .96 to 1; both syphilitic disorders, 2.75 to 1; all four organic disorders, 1.44 to 1; both functional disorders, 1.13 to 1; all six disorders, 1.22 to 1; all clinical groups, 1.09 to 1. It is thus seen that when the two races are equated as to general population (i.e., via the statistical device of standard rates of first admission), Georgia negroes are generally more likely than Georgia whites to be hospitalized with some form of mental disorder. Of the six specific psychoses studied, only in the case of senile psychoses do Georgia negroes have a lower probability of hospitalization than do Georgia whites. The chances of being hospitalized are markedly higher for negroes than for whites in the case of cerebral syphilitics (3.22 to 1), general paralytics (2.69 to 1), and schizophrenics (1.31 to 1).

For the benefit of the reader who may desire to compare negroes and whites with respect to the gross number of first admissions in each clinical category, disregarding the relative proportion of the two races in the general population, negro-white ratios of average annual first admissions are shown at the base of Table 3. These ratios may be interpreted as indicating the number of negro first admissions per 100 white first admissions. Using an average of the data furnished by the 1920 and 1930 Federal Census Reports, the general population of Georgia is found to be as follows: whites, 1,763,044; negroes, 1,138,745. That is, there were 65 negroes for every 100 whites in the general population of Georgia during the decade under study. When the negro-white ratio of the general population (.65 to 1) is compared with the negro-white ratios of the several clinical categories of the hospital population, it is found that negroes have higher than expected rates of first admission in ten of the twelve comparisons.

Comparisons of the percentages of first admissions and of deaths show marked variation from one category to another. When race is disregarded, the corresponding percentages of first admissions and of deaths, respectively, in the several clinical categories are as follows (Table 3, lines 4 and 5): senile, 3.7 per cent vs. 9.1 per cent; with cerebral arteriosclerosis, 9.1 per

cent vs. 14.6 per cent; general paralysis, 7.5 per cent vs. 13.6 per cent; with cerebral syphilis, .6 per cent vs. .8 per cent; manic-depressive, 25.2 per cent vs. 13.6 per cent; dementia praecox, 18.6 per cent vs. 13.7 per cent; both old-age disorders, 12.8 per cent vs. 23.7 per cent; both syphilitic disorders, 8.1 per cent vs. 14.4 per cent; all four organic disorders, 20.8 per cent vs. 38.1 per cent; both functional disorders, 44.1 per cent vs. 27.4 per cent; all six disorders, 64.9 per cent vs. 65.5 per cent. Although the six disorders combined account for an approximately equal percentage of all first admissions and of all deaths (64.9 per cent vs. 65.5 per cent), the percentage of total first admissions is markedly lower than the percentage of total deaths in each of the four organic disorders. In both of the functional disorders, this relationship is reversed. That is, manic-depressive and dementia praecox psychoses account for a higher percentage of all first admissions than of all deaths. Senile psychoses and manic-depressive psychoses constitute a lower percentage of all first admissions for negroes than for whites. The four remaining psychoses comprise a higher percentage of total negro first admissions than of total white first admissions.

The number of deaths per 100 resident patients constitutes an additional basis for comparing the several clinical groups among themselves, and of comparing one racial group with the other. In line 6 of Table 3, the following ratios of deaths per 100 resident patients are noted: senile, 34.8; with cerebral arteriosclerosis, 36.9; general paralysis, 38.5; with cerebral syphilis, 24.2; manic-depressive, 4.7; dementia praecox, 2.8; both old-age disorders, 36.1; both syphilitic disorders, 38.1; all four organic disorders, 36.8; both functional disorders, 3.5; all six disorders, 7.4; all clinical groups, 7.4. Here again it is observed that the six disorders combined have a death rate (7.4) identical with that for all clinical groups (7.4). At the same time, death rates for the four organic disorders are from three to five times as high as that for all clinical groups. The rate for dementia praecox (2.8) is 37.8 per cent of that for all clinical groups (7.4), and the rate for manic-depressive psychoses is only 63.5 per cent as high as that for all clinical groups. It is obvious

that the relatively low death rates found among these two functional disorders partly serve to explain why these disorders tend to constitute a higher percentage of the total resident population (58.1 per cent) than they do of total first admissions (44.1 per cent).

When the two races are compared with respect to deaths per 100 resident patients, the following negro-white ratios of standard death rates are obtained: senile, .62 to 1; with cerebral arteriosclerosis, 1.19 to 1; general paralysis, 2.80 to 1; with cerebral syphilis, .54 to 1; manic-depressive, 1.54 to 1; dementia praecox, 1.68 to 1; both old-age disorders, .95 to 1; both syphilitic disorders, 2.46 to 1; all four organic disorders, 1.36 to 1; both functional disorders, 1.64 to 1; all six disorders, 1.57 to 1; all clinical groups, 1.76 to 1. The above findings are considered to be of profound psychiatric, medical, and social significance. The obtained ratios indicate that, with minor exceptions (since senile psychoses and psychoses with cerebral syphilis constitute only 2.0 per cent and .6 per cent, respectively, of all negro resident patients), negroes have markedly or enormously higher standard death rates than do whites. To illustrate and interpret in the case of general paralytics: assuming the two races to have equal resident populations with this disorder, there are 28 negro deaths for each 10 white deaths. Using the same assumption as to equal resident populations for all clinical groups, there are seven negro deaths for each four white deaths.

Duration of Hospital Residence

Table 4 presents data dealing with the extent of hospital residence of patients dying during the decade. The amount and reliability of racial differences in hospital residence prior to death are shown in the lower section of Table 4.

Comparisons of disorders, disregarding race.—The median duration of hospital residence prior to death for all clinical groups is 1.46 years. Of the six distinct psychoses studied, only manic-depressive (2.20 years) and dementia praecox (7.20 years) psychoses show longer median periods of residence prior to death than that for all clinical groups. The median periods

TABLE 4

MEDIAN DURATION OF HOSPITAL RESIDENCE AT THE TIME OF DEATH OF PATIENTS IN EACH OF THE TWELVE CLINICAL GROUPS, BY RACE, FOR GEORGIA, 1923–32

Quartile Deviations of Distribution, Probable Errors of Medians, and Coefficients of Variation are also indicated. At the base comparisons are made of Whites and Negroes on the basis of median duration of hospital residence at the time of death. Critical ratios and chances in 100 of a true difference in medians are shown. All differences are positive (*i.e.*, in each comparison, Negroes have a shorter duration of hospital residence prior to death than do Whites). The numerals 1–12 under "Disorder Groups" refer to the following descriptive designations: (1) Senile; (2) Cerebral Arteriosclerosis; (3) General Paralysis; (4) Cerebral Syphilis; (5) Manic-Depressive; (6) Dementia Praecox; (7) Sum of Columns 1–2; (8) Sum of Columns 3–4; (9) Sum of Columns 1–4; (10) Sum of Columns 5–6; (11) Sum of Columns 1–6; (12) All Clinical Groups.

	Disorder Group (See legend)											
	1	2	3	4	5	6	7	8	9	10	11	12
Total												
Median	1.02	.66	.47	.70	2.20	7.20	.78	.48	.68	4.25	1.35	1.46
P.E. (dist.)	1.69	.82	.62	.63	3.07	7.87	1.00	.62	.83	5.50	2.17	2.37
P.E. (median)	.11	.04	.03	.27	.17	.43	.04	.03	.03	.23	.05	.05
V	85.22	83.67	75.61	78.06	73.44	75.39	81.38	74.70	78.95	82.22	86.14	89.52
Whites												
Median	1.31	.77	.78	.75	2.54	12.00	.95	.80	.91	6.45	1.71	2.00
P.E. (dist.)	1.85	.74	1.27	1.18	2.97	9.10	1.36	1.28	1.21	8.15	3.12	4.27
P.E. (median)	.14	.05	.12	.49	.23	.69	.07	.12	.05	.44	.11	.12
V	85.65	75.51	85.81	75.16	84.40	67.87	82.93	80.92	81.21	87.16	88.38	91.24
Negroes												
Median	.56	.50	.42	.67	2.00	4.80	.51	.44	.46	3.29	.99	1.06
P.E. (dist.)	.79	.76	.42	.61	3.76	3.91	.81	.42	.60	2.93	1.72	1.38
P.E. (median)	.10	.06	.03	.16	.29	.29	.05	.03	.03	.16	.06	.04
V	75.56	84.29	72.41	67.60	84.89	68.72	82.65	70.94	78.06	73.80	84.16	83.13
Negro-White Comparisons												
Differences in median	.75	.27	.36	.08	.54	7.20	.44	.36	.45	3.16	.72	.94
Critical ratios	4.27	4.29	2.98	.30	1.47	9.60	1.98	2.98	6.19	6.77	5.98	5.43
Chances in 100 of a true diff.	100	100	98	58	84	100	91	98	100	100	100	100

of residence prior to death for each of the twelve clinical categories are, in order of size, as follows: dementia praecox, 7.20 years; both functional disorders, 4.25 years; manic-depressive, 2.20 years; all clinical groups, 1.46 years; all six disorders, 1.35 years; senile, 1.02 years; both old-age disorders, .78 years; with cerebral syphilis, .70 years; all four organic disorders, .68 years; with cerebral arteriosclerosis, .66 years; both syphilitic disorders, .48 years; general paralysis, .47 years. It is evident that marked differences exist between the several groups in the extent of residence prior to death. Most of the obtained differences are statistically reliable, although data supporting this assertion have not been included in the tables. Of patients dying during hospital residence, those with dementia praecox live more than fifteen times as long as do those with general paralysis. In view of their comparative ages at the time of first admission (68.91 years and 62.58 years, respectively), it is worthy of note that patients with senile psychoses have a longer period of residence prior to death than do those with psychoses with cerebral arteriosclerosis (1.02 years and .66 years, respectively).

Quartile deviations of distribution and coefficients of variation respectively furnish absolute and relative measures of variability in the duration of hospital residence for each of the clinical categories. When quartile deviations are used as a criterion of variability in the duration of hospital residence prior to death, the following indices, stated in order of size, are noted: dementia praecox, 7.87 years; both functional disorders, 5.50 years; manic-depressive, 3.07 years; all clinical groups, 2.37 years; all six disorders, 2.17 years; senile, 1.69 years; both old-age disorders, 1.00 years; all four organic disorders, .83 years; with cerebral arteriosclerosis, .82 years; with cerebral syphilis, .63 years; both syphilitic disorders, .62 years; general paralysis, .62 years. Due to the fact that each of the distributions is markedly bunched toward the lower end of the curve, exact interpretation of these quartile deviations is difficult. It is evident, however, that the several clinical groups vary enormously among themselves with respect to absolute variability in the duration of hospital residence prior to death. For example, 75 per cent of the general paralytics

who die in residence do so within 1.09 years, whereas the corresponding figure for dementia praecox patients is 15.07 years.

The coefficients of variation, stated in order of size, are as follows: all clinical groups, 89.52; all six disorders, 86.14; senile, 85.22; with cerebral arteriosclerosis, 83.67; both functional disorders, 82.22; both old-age disorders, 81.38; all four organic disorders, 78.95; with cerebral syphilis, 78.06; general paralysis, 75.61; dementia praecox, 75.39; both syphilitic disorders, 74.70; manic-depressive, 73.44. When considered from this criterion of relative variability, the amount of difference between the several clinical categories is comparatively slight.

Inter-racial comparisons. — Negroes consistently die during much shorter periods of hospital residence than do whites. The amounts of difference between the two races in median duration of hospital residence prior to death among the several clinical categories are, stated in order of size, as follows: dementia praecox, 7.20 years; both functional disorders, 3.16 years; all clinical groups, .94 years; senile, .75 years; all six disorders, .72 years; manic-depresive, .54 years; all four organic disorders, .45 years; both old-age disorders, .44 years; both syphilitic disorders, .36 years; general paralysis, .36 years; with cerebral arteriosclerosis, .27 years; with cerebral syphilis, .08 years. The racial difference obtained in the case of patients with dementia praecox is enormous, and doubtless has more psychiatric, medical, and social significance than we shall attempt to evaluate here. In each of the twelve comparisons negroes have a markedly or appreciably shorter duration of hospital residence prior to death than do whites. Seven of the twelve differences have complete statistical reliability, and four of the remaining five differences have marked statistical reliability.

A comparison of the relative variability of the two races with respect to duration of hospital residence perhaps can best be made in terms of the negro-white ratios of coefficients of variations: *i.e.*, by dividing the twelve negro coefficients of variation by the corresponding white coefficients of variation. The ratios thus obtained are as follows: senile, .91 to 1; with cerebral arteriosclerosis, 1.12 to 1; general paralysis, .84 to 1; with

cerebral syphilis, .90 to 1; manic-depressive, 1.01 to 1; dementia praecox, 1.01 to 1; both old-age disorders, 1.00 to 1; both syphilitic disorders, .88 to 1; all four organic disorders, .96 to 1; both functional disorders, .85 to 1; all six disorders, .95 to 1; all clinical groups, .91 to 1. As judged by this criterion, it is seen that only in the case of psychoses with cerebral arteriosclerosis are negroes appreciably more variable than whites in duration of hospital residence. In most of the comparisons, negroes are slightly or significantly less variable than whites.

Age at Death

Comparisons of the several disorder groups among themselves and between the two races will now be made with respect to age at death. The upper portion of Table 5 provides medians, quartile deviations of distribution, P.E.'s of medians, and coefficients of variation for each of the clinical categories, by race. The amounts and reliabilities of racial differences in age at death are shown at the base of Table 5.

Comparisons of the disorders, disregarding race.—Marked variations exist among the several clinical groups in median ages at death. Senile psychoses (69.06 years) have the highest median age at death, while dementia praecox psychoses (43.28 years) have the lowest median age at death—the difference between the two median ages being 25.74 years. Stated in order of size, the median ages at death for the six specific psychoses are: senile, 69.06 years; with cerebral arteriosclerosis, 65.06 years; manic-depressive, 46.70 years; with cerebral syphilis, 44.17 years; general paralytics, 43.88 years; dementia praecox, 43.28 years. The six disorders combined (54.57 years) have a considerably higher median age at death than do all clinical groups (49.13 years).

The quartile deviations of distribution also show considerable variation from one disorder group to another. Stated in order of size, the quartile deviations for each of the several clinical groups are as follows: all clinical groups, 15.10 years; all six disorders, 14.12 years; manic-depressive, 13.82 years; both functional disorders, 13.76 years; dementia praecox, 13.65 years; with

TABLE 5

Median Ages at the Time of Death of Patients in Each of the Twelve Clinicacl Groups, by Race, for Georgia, 1923–32

Quartile Deviations of Distribution, Probable Errors of Medians, and Coefficients of Variation are also indicated. At the base comparisons are made of Whites and Negroes on the basis of age at the time of death. Critical ratios and chances in 100 of a true difference in medians are shown. All differences are positive (*i.e.*, in each comparison, Negroes die at an earlier age than do Whites). The numerals 1–12 under "Disorder Groups" refer to the following descriptive designations: (1) Senile; (2) Cerebral Arteriosclerosis; (3) General Paralysis; (4) Cerebral Syphilis; (5) Manic-Depressive; (6) Dementia Praecox; (7) Sum of Columns 1–2; (8) Sum of Columns 3–4; (9) Sum of Columns 1–4; (10) Sum of Columns 5–6; (11) Sum of Columns 1–6); (12) All Clinical Groups.

	Disorder Group (See legend)											
	1	2	3	4	5	6	7	8	9	10	11	12
Total												
Median	69.06	65.06	43.88	44.17	46.70	43.28	67.06	43.92	60.90	45.00	54.57	49.13
P.E. (dist.)	5.16	8.10	8.71	13.00	13.82	13.65	6.47	8.94	12.01	13.76	14.12	15.10
P.E. (median)	.35	.42	.47	2.65	.74	.75	.27	.44	.39	.53	.35	.30
V	7.47	12.58	19.67	28.57	29.16	30.59	9.65	20.11	20.55	29.92	26.27	30.47
Whites												
Median	69.72	67.17	45.55	51.67	56.32	50.67	68.46	45.80	65.41	53.68	62.26	56.39
P.E. (dist.)	5.14	6.41	8.79	8.96	13.19	13.65	5.78	8.87	8.66	13.41	11.79	14.28
P.E. (median)	.40	.47	.82	3.36	1.02	1.05	.71	.86	.39	.73	.41	.40
V	7.37	9.55	19.11	17.85	23.88	26.41	8.44	19.25	13.51	25.01	17.00	26.24
Negroes												
Median	68.18	61.08	43.06	40.00	38.79	36.33	63.68	42.91	52.84	37.52	46.30	42.57
P.E. (dist.)	6.36	8.95	8.42	13.34	10.98	10.38	8.63	8.65	12.28	10.73	12.74	12.99
P.E. (median)	.79	.71	.57	3.57	.85	.81	.59	.64	.58	.59	.45	.38
V	9.33	14.55	19.43	30.09	27.71	27.40	13.64	19.97	23.02	27.69	27.01	29.79
Negro-White Comparisons												
Differences in medians	1.54	6.09	2.49	11.67	17.53	14.34	4.78	2.89	12.57	16.16	15.96	13.82
Critical ratios	1.67	7.10	2.50	2.40	13.28	10.80	5.10	2.93	17.90	17.23	26.00	25.10
Chances in 100 of a true diff.	87	100	95	95	100	100	100	98	100	100	100	100

cerebral syphilis, 13.00 years; all four organic disorders, 12.01 years; both syphilitic disorders, 8.94 years; general paralysis, 8.71 years; with cerebral arteriosclerosis, 8.10 years; both old-age disorders, 6.47 years; senile, 5.16 years. It is thus seen, for example, that while 50 per cent of senile deaths occur between the ages of 63.90 years and 74.22 years, the corresponding scatter in the case of all clinical groups is from 34.03 years to 64.23 years.

That the relative variability in age at death also differs from one disorder group to another is seen from the following coefficients of variation (stated in order of size): dementia praecox, 30.59; all clinical groups, 30.47; both functional disorders, 29.92; manic-depressive, 29.16; with cerebral syphilis, 28.57; all four organic disorders, 20.55; both syphilitic disorders, 20.11; general paralysis, 19.67; with cerebral arteriosclerosis, 12.58; both old-age disorders, 9.65; senile, 7.47. From these comparisons it is seen, for example, that seniles are roughly only one-fourth as variable in age at death as are all clinical groups.

Inter-racial comparisons.—It is seen from Table 5 that in each of the twelve clinical groups negroes die at an appreciably or markedly earlier age than do whites. The amounts of racial difference in median age at death vary from one clinical category to another, to wit: manic-depressive, 17.53 years; both functional disorders, 16.16 years; all six disorders, 15.96 years; dementia praecox, 14.34 years; all clinical groups, 13.82 years; all four organic disorders, 12.57 years; with cerebral syphilis, 11.67 years; with cerebral arteriosclerosis, 6.09 years; both old-age disorders, 4.78 years; both syphilitic disorders, 2.89 years; general paralysis, 2.49 years; senile, 1.54 years. The racial difference in median ages at death is more than eleven times as great for manic-depressive psychoses (17.53 years) as for senile psychoses (1.54 years). Eight out of the twelve obtained differences have complete statistical reliability and the remaining differences are sufficiently reliable to suggest the probability of true racial differences in age at death.

The negro-white ratios of coefficients of variation reveal the presence of racial differences in relative variability in age at

death. The obtained negro-white ratios of coefficients of variation for each of the several clinical groups, stated in order of size, are as follows: all four organic disorders, 1.70 to 1; with cerebral syphilis, 1.69 to 1; both old-age disorders, 1.62 to 1; all six disorders, 1.59 to 1; with cerebral arteriosclerosis, 1.52 to 1; all clinical groups, 1.31 to 1; senile, 1.27 to 1; manic-depresive, 1.16 to 1; both functional disorders, 1.11 to 1; both syphilitic disorders, 1.04 to 1; dementia praecox, 1.04 to 1; general paralysis, 1.02 to 1. It is seen from the above ratios that negroes are more variable than whites with respect to age at death. Stated differently, negroes are generally less alike among themselves in age at death than are whites.

Summary and Conclusions

The data presented for the twelve clinical categories of white and negro patients in Georgia during the decade 1923–32 appear to warrant the following conclusions:

(1) Marked differences exist among the several clinical groups in median age at first admission, and in absolute (P.E. of distribution) and relative (coefficient of variation) variability in age at first admission.

(2) With the exception of psychoses with cerebral syphilis, negroes have an appreciably or markedly lower median age at first admission than whites in every comparison made, and most of the obtained racial differences have marked or complete statistical reliability.

(3) A study of the distribution of first admissions among the several clinical categories, by age levels, indicates that as one passes from younger to older age groups, the psychoses vary in their relative frequencies in characteristic fashion. At the youngest age intervals, manic-depressive psychoses constitute, for both races, a higher percentage of all patients than does any other type of psychosis. The middle age intervals are characterized by relatively high rates of first admissions among general paralytics, manic-depressives and schizophrenics. In the older age intervals, psychoses with cerebral arteriosclerosis, and senile psy-

choses predominate. There is a generally greater tendency among negroes than among whites for the disorders under study to constitute a larger proportion of all first admissions, particularly at the earlier age levels.

(4) The standard rates of first admission vary markedly from one clinical category to another. Of the six specific psychoses studied, the standard rate for manic-depressive psychoses is the highest both for whites and for negroes. Both races have relatively low standard rates among senile psychoses and psychoses with cerebral syphilis.

(5) The negro-white ratios of gross first admissions vary strikingly from one clinical category to another. For example, there are only 43 senile first admissions among negroes for each 100 senile first admissions among whites, whereas for each 100 white first admissions with cerebral syphilis there are 220 negro first admissions with cerebral syphilis. As contrasted with these findings, there were 65 negroes for each 100 whites in the general population of Georgia during the decade under study.

(6) Although Georgia negroes have a smaller gross number of first admissions than Georgia whites in most of the comparisons herein reported, the standard rates of first admissions for negroes are appreciably or markedly higher than the corresponding standard rates for whites, with the single exception of senile psychoses. Comparisons of the standard rates of first admission for the two races indicate that the general probability of hospitalization is markedly higher for negroes than for whites, particularly among cerebral syphilitics (3.22 to 1), general paralytics (2.69 to 1), and schizophrenics (1.31 to 1).

(7) The negro-white ratios of deaths per 100 resident patients indicate that, with minor exceptions, negroes have markedly or enormously higher standard death rates than do whites. For each four white deaths among all white hospital patients there are seven corresponding negro deaths; for each 10 white deaths among general paralytics there are 28 corresponding negro deaths.

(8) The duration of hospital residence prior to death varies markedly from one disorder to another as is evidenced by the

fact that dementia praecox patients have a median period of hospital residence prior to death more than fifteen times as great as that for general paralytics.

(9) In each of twelve comparisons negroes are found to have appreciably or markedly shorter periods of residence prior to death than whites, and most of the obtained racial differences are statistically reliable.

(10) The median age at death varies enormously from one disorder to another. Of the disorders studied, senile psychoses (69.06 years) have the highest median age at death while dementia praecox psychoses (43.28 years) have the lowest median age at death.

(11) Negroes are found to die at a much earlier age than whites in each comparison made. The greatest racial difference occurs among manic-depressive psychoses (17.53 years), while the smallest difference is found among senile psychoses (1.54 years). All of the obtained differences have marked or complete statistical reliability.

The findings herein reported, incomplete and imperfect as they admittedly are, appear to the writer to have considerable psychiatric and social significance. However, it is suggested that the chief significance of this study lies in the fact that it has revealed research leads relative to marked racial differences in mental disorders which are worthy of further detailed study. It appears that in Georgia, negroes are more likely than whites to become insane, particularly with syphilitic disorders and dementia praecox. Also, as compared with Georgia whites, Georgia negroes have notably high death rates. Georgia negroes die during much shorter periods of hospital residence and at much earlier ages than Georgia whites. It should be profitable to compare these findings for Georgia with conditions in other areas. Further studies should be made to determine as exactly as possible the causes of the conditions herein reported. The present data do not justify the assumption of a constitutional inferiority among negroes, nor do they prove that the differences are due to racial discrimination or exploitation.

THE MEASUREMENT OF ATTITUDES TOWARDS MATHEMATICS [1]

EURI BELLE BOLTON
GEORGIA STATE COLLEGE FOR WOMEN

THE CONCEPT AND MEASUREMENT OF ATTITUDE IN SOCIAL PSYCHOLOGY

During the past three quarters of a century psychologists and sociologists have tried to develop an acceptable theory of attitudes and to determine their relation to behavior, with the result that the concept has become one of the most important in social psychology. It is only recently, however, that a sufficient degree of uniformity in the comprehension and use of the term has been arrived at to insure progress in the analysis of attitudes.

1. *Modern conceptions of attitudes.* The conclusions from the earlier attitude studies of the modern period were indefinite and conflicting because the usages of the term were varied, and the questionnaire and life history methods first used in attitude measurement were inaccurate. Bain (**4**, **5**) in 1927 directed attention to the unscientific use of the term in sociology. His conception that an attitude is the behavior itself and not a preparation for behavior has not been accepted generally either by sociologists or social psychologists, but his critical analysis

[1] The writer wishes to express appreciation to the following persons who have generously coöperated in making this study possible: to Dr. Edwards of the University of Georgia and to Dean Henderson of the South Georgia Teachers College, for securing statements of attitudes from students at these institutions; to Miss Martin and Miss Hallie Smith for the coöperation of their students in giving and judging statements of attitudes; to Mr. Knox and Mr. Massey of G. S. C. W., and to Dr. Wren and the members of the Mathematics Department of George Peabody College for Teachers, for permitting students in their classes to take the attitude tests; and especially to Miss Napier, Miss Nelson and Dr. Walden of the Mathematics Department of G. S. C. W. for criticizing and judging the attitude statements. Dr. Walden has read the manuscript and has given many helpful criticisms. The writer is also indebted to Dr. H. H. Remmers of Purdue University for permission to use the Purdue Generalized Scale to Measure Attitude Toward Any School Subject.

showed that progress could not be made in attitude study until the concept was clarified and the problem attacked by the techniques of controlled experiment. Thurstone (45, 46) in 1928, influenced by the earlier study of F. H. Allport and Hartman showing that opinions from a large number of students on public issues like prohibition and the League of Nations tended to distribute themselves according to the normal frequency curve, applied the psycho-physical method of equal appearing intervals to the construction of attitude scales. He assumed that attitudes may be conceived as constituting a linear continuum which expresses differences in degree of affect for or against a social object or issue and that verbal statements of opinions which express different degrees of the attitude variable may be determined by finding the average comparative values assigned to them by a large number of judges who arrange them on a scale according to their discriminable differences. The formulation of this concept made possible for the first time the application of a reliable statistical procedure to the analysis of attitudes.

Droba (15, 16, 17, 18, 19) has attacked the problems of the concept of attitude both by experiment and theoretical analysis. He defined an attitude as . . . "a mental disposition to act for or against a definite object," and thinks that it results from experiences . . . "that have been molded into a totality that is too complex and too intricate to understand" (17, p. 451). He distinguishes between attitude and motive, saying that a motive starts behavior and an attitude determines its direction, and emphasizes the importance of clearly defining the objects of reference towards which attitudes are directed. At present practically nothing is known about the neural basis of attitudes, but Lashley (30, p. 23) has shown that this is also true of learning and intelligence. A number of studies have shown, however, that it is possible to analyze the causal factors resulting in the development and change of attitudes. G. W. Allport (3) concludes from a review of the studies of the genesis of attitudes that they are acquired by the integration of numerous responses, of a similar type, by individuation or segregation, by traumatic experiences and by the process of adopting ready-made the atti-

tudes of parents, teachers, and playmates. The present evidence seems to indicate that attitudes are learned just as perceptual knowledge and habits are acquired and, therefore, that the explanation of the neural basis of attitudes will have to be determined by the same concepts and procedures as those used in studying learning. Lashley (30) and Peterson (34) emphasize the necessity of formulating any theory of learning and brain activity in terms of the whole organism and not of elementary reflexes. Any concept of attitudes which considers their neural basis would undoubtedly have to conform to the same view. The exact relationship of attitudes to behavior must yet be determined by further experimentation, but observation and *a priori* analysis indicate that attitudes may motivate behavior. Katz and Allport (27, p. 369) and G. W. Allport (3) think that this is true. The latter distinguishes, however, between highly energized attitudes that drive or motivate behavior and those that are merely directive.

Three of the most important recent contributions to the clarification of the concept of attitude have been made by Cantril, by Likert and G. W. Allport. Cantril (9) subjected the theory of the specificity of attitudes to experimental analysis and his data show that attitudes are generalized, that they are more constant and enduring than specific content and that a general attitude seems to serve as a dynamic or directive influence on more specific attitudes and reactions.

Likert's study (31) of attitudes towards internationalism, prohibition and the Negro, also, gives evidence that attitudes are general rather than specific. He suggests, however, that it may be possible to discover two or more attitude variables that are closely related and directed towards the same social object or issue. The writer in an experimental analysis of the Hinckley Scale of Attitudes toward the Negro (6, 7) obtained data which led to the tentative conclusion that there may be an attitude variable toward social intermixture with the Negro race which differs in individuals and in groups from the attitude towards the social rights of the Negro in regard to cultural, economic and political advantages.

G. W. Allport (3) has traced the historical development of the concept of attitude from the time it was first used by Herbert Spencer in 1862 and has analyzed the more recent studies of attitude. He distinguishes attitude from the following types of readiness for response: reflexes; conditioned reflexes; habits; instincts or innate tendencies to action; needs, wishes or desires; sentiments; concepts and traits. He says there is a good deal of difference of viewpoint in regard to the relation of motor set to attitude and in regard to whether or not verbal opinion is an expression of an attitude. His criticism of the definitions usually given is that they do not distinguish between attitudes which are often very general and habits which are always limited in their scope. His definition includes, he says, the following recognized types of attitudes: the *aufgabe,* the quasi-need or temporary attitude, the *Bewusstseinlage,* interest and subjective value, prejudice, stereotype and even the broadest conception of all, the philosophy of life. It excludes those types of readiness that are expressly innate, that are bound rigidly and invariably to the stimulus and those that lack flexibility, directionality and reference to some external or conceptual object. The specific definition is: *"An attitude is a mental state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situation with which it is related"* (3, p. 810). The necessity for giving such a detailed explanation of the definition shows the great confusion in the general use of the term. It undoubtedly indicates, also, that any theoretical concept of attitudes formulated on the basis of present knowledge must be subjected to further rigid experimentation.

2. *Techniques for measuring attitudes.* Critical summaries of the methods used in measuring attitudes have been given by Bain (5), Droba (15), Katz and Allport (27), Fryer (21), and Murphy and Murphy (33). Only a brief reference to this phase of the problem will be made here.

The most significant contrbiution to the development of statistical techniques for the measurement of attitudes has been made by Thurstone (44, 45, 46, 47, 48). He has applied the

method of equal-appearing intervals to the construction of scales of opinions in regard to social issues or values. By securing a large number of statements of opinion in regard to certain values from many individuals in different groups and having them judged by a number of judges to determine the degree of favorableness or unfavorableness which they express, it is possible to choose from the original list a small number of valid statements which constitute a linear continuum varying from attitudes that are extremely negative to those that are extremely positive in regard to the value. When the first scales were constructed by Thurstone and his students, it was assumed that the scale values of statements are not affected by the individual attitudes of the judges. This assumption has been verified by Hinckley (24), by Ferguson (20) and more recently by Pintner and Forlano (35).

The most important criticisms of the Thurstone method may be summarized briefly as follows: It has been considered too time-consuming. Seashore and Hevner (40) have modified the method by having the judges keep the scale intervals before them and assign the scale values to the statements arranged on a mimeographed sheet instead of having them on cards and arranged into piles. This method lessens considerably the time needed for the preparation of the statements and, also, that necessary for judging them.

Likert (31) has applied the sigma scoring technique to determine the scale value of items included in his attitude scales previously referred to. He discovered empirically that the responses to statements having five steps for expressing agreement or disagreement tended to be distributed according to the normal frequency curve and he assigned sigma values to represent the percentage of persons who checked each of the five degrees of approval or disapproval for each statement. The sigma technique was further simplified by assigning the numbers *one, two, three, four* and *five* instead of sigma values to the positive and negative values on the scale. *One* was always applied to the negative end of the sigma scale and *five* to the positive end. Scoring by the simpler method yielded practically the same

results and Likert thinks is better because it makes possible such a tremendous saving in time.

Williamson and Darley (50, 51) have verified the findings of Likert in the construction of a social preferences scale and a social adjustment scale. Each of the scales includes forty statements and each statement is worded so that it is possible to express agreement or disagreement in terms of five degrees of approval or disapproval. The coefficients between the tests when scored arbitrarily by the simple method of numerical weights and the sigma method varied from .94 to .98 and the simple scoring method was adopted. Rundquist and Sletto (39) used the Likert method in the construction of five scales for measuring the effects of unemployment on the personality traits and attitudes of unemployed groups. They found the arbitrary method of scoring to be as good as the sigma method.

Though he considers the Thurstone method of measuring attitudes the most important contribution yet made to attitude measurement, G. W. Allport (3) questions the assumptions on which the method is based. He says that attitudes are not necessarily arranged on a single continuum, but are often discrete and highly individual. There is, also, the question of whether the scale values for statements derived from one population are applicable to other populations of subjects. Likert's results show, he thinks, that for many studies the logical *a priori* scale is more practicable. Guilford (23, p. 163) thinks that the validity of the application of the method of equated psychological intervals to attitude measurement has not been adequately established. Katz suggests that it would be more practicable to use the *a priori* type scale to discover attitudes and their relation to possible alternative attitudes, and when they have been . . . " brought out by this method the affective intensity of each attitude can be measured by the psycho-physical scale " (28, p. 480). Until more comparative results from the use of the various scaling techniques are available, this would be desirable. When comparable scales are developed for investigating the intensity of attitudes that are theoretically very closely related, the Thurstone technique would seem to be preferable.

3. *Factors influencing changes in attitudes.* One of the most general applications of attitude scales has been the measurement of changes in attitude due to environmental influences or controlled educational experiences. Thurstone (47) found that seeing a moving picture film once resulted in a reliable difference in the attitudes of children toward the Chinese. He and Peterson have made other studies which show that motion picture films affect the attitudes of children.

Droba (16) found that juniors and seniors in regular college classes were more liberal in their attitudes towards the Negro than freshmen in the same college groups. Sowards (41) found that college freshmen were more opposed to war than high school seniors and that college seniors were more opposed to war than freshmen. Her study indicates that education tends to develop more decided attitudes for peace. Moore and Garrison (32) found a slight though general tendency for reactionary and conservative attitudes towards questions dealing with problems of politics, sex, family, individual freedom, economic relations, authority of religious and national tradition and international and racial relations to decrease from the freshman to the senior class in college. A comparison of a limited number of cases whose scholarship records were available seemed to suggest that those whose scholarship record is lowest tend to have either more reactionary or more conservative views.

Acheson (1) studied the relation of age to a liberal attitude towards modern social problems. Women graduate students from thirty to thirty-four years of age were more liberal and flexible in their attitudes towards such problems as reading scientific sex literature, moderate smoking for women, paying one's own way on a date, divorce, moderate drinking for men, etc., than women who were from forty to fifty-two years of age. Varied experiences, wide reading and graduate work in a liberal university were the most important influences causing changes in attitude toward these issues.

Carlson (10) found that the majority of a group of undergraduate seniors at the University of Chicago were opposed to prohibition, neutral in their belief in God and their attitude

toward communism and sympathetic toward pacifism and birth control. The men were more opposed to prohibition and believed less in the reality of God, but there was no difference between the attitudes of men and women towards the other issues. Students majoring in the physical sciences were more conservative in their attitudes than those majoring in the social sciences. There was no relation between intelligence and attitudes towards prohibition, but the more intelligent were more liberal in their attitudes towards the other issues. Early religious training seemed to be a factor in determining the attitudes of the group. The results obtained indicate that the attitudes of this group had been influenced by intelligence, by religious training and by an individual tendency to be radical or conservative.

Knower (29) obtained results that show a reliable change in the attitudes of subjects after they heard an oral argument for or against prohibition. Control groups of subjects were used. The change in the attitudes of subjects in the experimental groups was five times the change in the control groups. Logical and persuasive speeches were equally effective in producing changes in attitude. It was more effective to present the argument to a subject in a room alone than to present it to a group of subjects in the audience situation. Women changed their attitudes in greater numbers and to a greater extent than men. Men and women speakers secured attitude changes that were approximately equal, but women were more influenced by men speakers and men were more influenced by women speakers.

Telford (42) investigated the effects of studying courses in criminology, in sociology, in educational psychology and in general psychology on attitudes towards the treatment of criminals. The change in attitude resulting from the courses was greatest in criminology and was least in general psychology. Those whose attitudes towards the problem were extreme and those who were more intelligent were more consistent in checking the statements. Results obtained in the same study indicate that denominational teaching is a factor in determining loyalty to the church. The denominational groups were loyal to the church in the following order: Catholics, Baptists, Lutherans, Methodists,

Presbyterians, Episcopalians and Congregationalists. Cherrington (12) studied the effects of a lecture and of reading articles opposing war upon the attitudes of freshman and sophomore students toward war. The differences in attitude resulting from the reading and the lecture were reliable and persisted when measured again six months later. Gardner (22) found that lectures, stories and chalk talks resulted in a reliable change of attitudes toward prohibition and war. Pupils in the junior high school were more influenced by the stimuli applied than were college freshmen.

One of the most elaborate studies of the influence of social stimuli on attitudes has been made by Remmers and his students (36, 37). They have obtained results which show the following effects of social stimuli on the attitudes of different groups of junior and senior high school and college students: that student government caused junior high school students to have a more favorable attitude towards law observance; that a critical study of poetry resulted in a reliably greater appreciation for poetry; that reading literature highly favorable towards the Negro resulted in a more favorable attitude toward the Negro race; and that hearing a single lecture favorable to the League of Nations resulted in a change to a more favorable attitude toward it. They have shown, also, that it is possible to secure reliable changes in attitudes toward social issues in any direction desired by presenting stimulus material in the form of lectures or papers prepared expressly for that purpose.

All of the investigations of the factors causing changes in attitudes towards social issues reviewed were made by having subjects check verbal statements of opinion which represent the attitudes. Katz and Allport (27) have pointed out the possible inaccuracies in results obtained by this method because people seem to have private and public attitudes toward certain social issues. Corey (13) has studied the relationship between expressed verbal opinions in regard to honesty and actual cheating on examinations. His measure of cheating was the difference between the true scores obtained on examinations in educational psychology and scores assigned by the students when

permitted to grade their own papers. He obtained a correlation of .024 $\pm$ .12 between the actual cheating scores and those made on a test of attitude toward honesty constructed by the Thurstone technique. Corey concludes that a person's statements of verbal opinion about honesty have no value in predicting his practice of cheating and suggests the necessity for caution in accepting the validity of studies which attempt to measure the influence of social stimuli on changes in attitude. Cantril (9) classifies honesty as an ethical concept and not as a common attitude. It is possible that permitting the students to correct their own papers over a period of time without any reference to cheating may have caused them to acquire a temporary attitude toward the course which affected their cheating behavior. F. H. Allport (2) has recently pointed out that personality traits must be measured in terms of what the person is trying to do and must be interpreted in relation to the situation in which the behavior occurs. Corey's study and conclusions show, however, the complexity of the problem of measuring both personality traits and attitudes. Behavior scales of the type recently developed by Rosander (38) may prove to give a better measure of those attitudes which motivate behavior than the opinion scales which have been used in the studies previously developed, but this application of the Thurstone technique to scale construction will have to be verified by much further experimentation before its value is established.

The Problem of the Present Study

One of the most extensively investigated phases of attitude measurement is that of the analysis of the development of interests and their relation to behavior. Fryer (21) has given a very complete summary and criticism of the studies in this field. He distinguishes between vocational, educational, and social and intellectual interests. He concludes from the studies analyzed that vocational interests are genetic in development, but are relatively so unstable that measurement of them in childhood and adolescence has little value for prediction of future interests or for practical purposes of vocational guidance. The total interest picture is more valuable for understanding the adjust-

ment of the individual. Educational interests seem to have very little value for prediction of achievement and scholarship. There are, he says, common group interests and they exist in greater numbers among occupational than among ability groups.

Caswell and Campbell (11), who accept Fryer's conception of interests, are of the opinion that the great volume of attitude studies appearing in psychology in recent years have as yet had practically no effect on the school curriculum. As a result of what psychology has thus far accomplished in attitude measurement, however, they and other educational leaders are including desirable attitudes and interests as aims or goals for determining the materials of instruction which will contribute to the development of the whole personality of the child.

Katz (28) has recently suggested that attitude measurement heretofore has been largely practical and that it must be applied more to the problems of social psychology if it is to survive. An analysis of the studies reviewed above shows that even in the practical fields of applied psychology the results do not yet assure the survival of this phase of the movement.

Remmers and his students (36, 37) have attacked the problem of the relation of attitude measurement to the school curriculum in their studies of attitudes towards school subjects. The interests of children of different ages and the factors affecting the development of individual interests should be analyzed in order to aid in determining the materials and activities of the curriculum suitable for children of different age levels and abilities, but, since attitudes are important motivating factors in behavior, it is equally as important from the standpoint of education in relation to society and human welfare to determine what attitudes result from definitely controlled stimuli in the formal school program.

Silance (36) working under Remmers' direction has constructed a generalized scale for measuring attitudes towards any school subject. Remmers, Taylor and Kintner (37) used the scale to measure the attitudes of freshmen at Purdue towards English, Biology, Chemistry and Mathematics. The test was given at Thanksgiving during the Fall Quarter. The students

had a favorable attitude towards all the subjects involved, but those majoring in the different schools of the university differed in their attitudes toward the various subjects. Many of the differences obtained were not reliable and merely indicate a general trend. The most important sex differences were that the women favored English more and mathematics and chemistry less than the men.

There have been conflicting opinions about the cultural and practical values of mathematics since Protagoras advanced his view (**8**). At various periods in the progress of education conflicting theories in regard to the values of mathematics have appeared. A renewed effort to determine the functional and social values of mathematics is now being made both by educational philosophers and mathematicians. Dewey said in 1916 in explaining his conception of interest, " Numbers are not objects of study just because they are numbers already constituting a branch of learning called mathematics, but because they represent qualities and relations of the world in which our action goes on, because they are factors upon which the accomplishment of our purpose depends. . . . This connection of an object and a topic with the promotion of an activity having a purpose is the first and last word of a genuine theory of interest in education " (**14**, p. 158). He says further in regard to the relation of experience to thinking that though the thinking act is born in " partiality " or in a personal sharing in what is going on it must in order to achieve its tasks become " impartial " or impersonal. " Only gradually," he says, " and with a widening of the area of vision through a growth of social sympathies does thinking develop to include what lies beyond our *direct* interests: a fact of great significance for education " (**14**, p. 173). He does not seem to make any distinction between an interest in a subject and its intrinsic value, but does distinguish between interest and an instrumental value. If the subject matter of the curriculum, which is of necessity highly symbolical, is used to enrich the experience of those who study it and they are actively interested in the intellectual progress made, the result is a development of appreciation of values. If the interest is in the experience itself,

the values of the experience are intrinsic and are not measurable; if, however, the experience is interesting because it is the means of realizing some other desirable purpose or purposes, the values are instrumental and are measurable. If the number relations and processes of mathematics are taught as Dewey has suggested as a means of interpreting the social relationships which they represent, then it would seem that the intellectual growth experienced would result in the development of attitudes of appreciation that would be of great value to the individual in interpreting the organized group life of which he is a part. Judd (26) has illustrated this conception of number relations in his explanation of the significance of the measurement of time to the coöperative group activities that characterize modern civilized life.

If the writer has interpreted Dewey's view correctly, it is similar to G. W. Allport's conception of the relation of interests to values. Allport says, " Interests are a special type of enduring attitudes which refer, as a rule, to a class of objects rather than to a single object. They are dynamic attitudes, rich in ideational content, and involve a recognition and understanding of the objects which have satisfying properties. Subjective values are essentially the same as interests, except they are more properly spoken of when the individual is mature and has reflected upon and organized his interests within a comprehensive and consistent system of thought and feeling " (3, p. 809). It was pointed out above that Fryer concluded from his analysis of the investigations of interests that there are common interests which vary in different groups. It is the purpose of the present study to develop two comparable attitude scales, by the Thurstone-Chave method of equal-appearing intervals, to be used for an empirical investigation of the possible existence of two related common attitudes of interest in and appreciation of the value of mathematics.

Construction of the Scales

Students at the Georgia State College for Women, at the South Georgia Teachers College and at the University of Georgia were

asked to write a statement of their attitudes towards mathematics. There were statements from 105 women at G. S. C. W., from 132 men and 192 women at S. G. T. C., and from 62 men and 28 women at the University. After the general statement had been written, the students were asked to write their opinion in regard to: (1) their interest in mathematics, (2) the values of mathematics, (3) the methods of teaching mathematics, in comparison with other school subjects. The last statement was asked for in order to secure empirical evidence of the relation between one's experience with mathematics and his attitudes toward the subject.[2] Additional statements were obtained from the literature dealing with the teaching of mathematics.

From the above sources, a list of 399 statements about mathematics was made up. The first list of 129 statements obtained from the G. S. C. W. students, after being edited by the writer, were evaluated independently by three teachers of mathematics in this institution. All statements considered doubtful or ambiguous or that gave any implication of the traditional formal discipline conception of mathematics were eliminated. The statements from the students in the other institutions came during the spring vacation and it was not possible to have them evaluated by the mathematics teachers. The original list of statements was reduced to two-hundred-twenty-six. The number of statements of interest and of value was approximately equal. Some statements in regard to the teaching of mathematics were included. They were as reliable statistically as the other statements, but they were not used in the construction of the final scales. The statements were mimeographed, mounted on small cards and arranged in sixteen duplicate sets.

The statements were judged by the Thurstone–Chave (44) method by 124 students in the freshman, sophomore, junior and senior classes at G.S.C.W. They were all taking courses either with the writer, with Miss Martin in Physics or with Dr. Walden

[2] The instructions given to the students at the South Georgia Teachers College were mimeographed instead of being read orally, but a comparison of the statements from this group with those from the other institutions showed that the variation in method did not affect the results.

in Mathematics. With the exception of one group, under the supervision of Miss Martin, the students worked under the writer's direction. Two of the methematics teachers judged the statements and the mean of their judgments was used as an empirical check on the ambiguity of the statements. If both of the mathematics teachers assigned a value to a statement that differed by more than two points from the median value of the students' judgments, it was thought that it might be interpreted differently by those who have different training in mathematics and was not used. A few undesirable statements were discovered by this method of checking. For example, one statement judged was " My limited success in mathematics has always worried me." The student giving this statement meant to express an unfavorable attitude, but both teachers of mathematics interpreted it to mean that the student was interested in mathematics and wanted to achieve more in the field and assigned it to a favorable value. The independent rating by two specialists in the field of mathematics served as a valuable check on the clarity of the statements.

The scale values were determined graphically by the Thurstone–Chave method (44) and those statements having high Q-values were not used in the scales. The Q-values for the statements included in the scale are given in Appendix I. No check of internal consistency was applied to the statements. From the 226 statements judged four comparable scales were set up, two for ' interests ' and two for ' values.' Each form originally included twenty-one statements, but two of the statements in the values scales were not valid and were eliminated from the final scales.

Reliability and Validity of the Scales

1. *Reliability.* Forms A and B of the Interest Scale, Forms A and B of the Values Scale and Form A of the Purdue Scale for measuring attitudes towards any school subject were mimeographed and clipped together. These were given at one sitting of about forty minutes of working time to 124 students who were attending the second session of the G. S. C. W. Summer School and to 35 students who were members of advanced

graduate classes in mathematics at George Peabody College for Teachers.

Five of the G. S. C. W. students and 28 of the Peabody students were majors in the field of mathematics. The Peabody students were all working either for the M.A. or the Ph.D. degree and most of them were teachers of mathematics. One of the group, who was an observer in one of the classes, had already received the Ph.D. degree and was a teacher of mathematics in an undergraduate college.

The reliability coefficients between Forms A and B of the Interest and the Values tests are based on all of the 159 cases included in the study. The uncorrected coefficient for the Interest test is .886 ± .012 and for the Values test it is .814 ± .019.

TABLE 1

RELIABILITY COEFFICIENTS OF THE ATTITUDE TESTS

Test	Number of Cases	r	P.E.
Interest—Form A and Form B...........	159	.886	.012
Values—Form A and Form B............	159	.814	.019

The reliability coefficient for the Purdue test was not computed. Remmers, Taylor and Kintner report a coefficient of .741±.013 for forms A and B of the scale, when it was given to 579 students as a measure of attitudes towards mathematics. The reliability of the Interest and the Values tests is quite satisfactory in comparison with that of other attitude scales. The reliability of the Values scale should be improved, however, if it is used for individual diagnosis.

2. *Validity of the Scales.* One method for testing the validity of an attitude scale is to determine the extent to which it differentiates between groups whose attitudes might seem on the basis of *a priori* judgment to differ decidedly in the degree of favorableness or unfavorableness toward the attitude object. This method of validating attitude scales was used by Williamson and Darley (49, 50), and by Rosander (38) and was one of the methods used by Likert (31) and by Rundquist and Sletto (39). It would seem that students who plan to teach a subject or who have taught it and are doing advanced graduate study in the field

of their choice would be much more favorable in their attitudes towards the subject than students majoring in other fields. The mean scores made by the G. S. C. W. and Peabody mathematics majors were compared with the mean scores made by students majoring in other fields. The differences and reliabilities of the differences in the mean scores made by the two groups on all of the tests are reliable, although the difference as measured by our Interest scale is greatest. It is not so great for values because some students majoring in other fields have a neutral interest

TABLE 2

Comparison of Mean Scores Made on the Attitude Towards Mathematics Tests by Those Majoring in Mathematics and Those Majoring in Other Subjects

Group	N	Interest A	Interest B	Values A	Values B	Interest A & B	Values A & B	Purdue A
Math. Majors	33	7.198	7.76	7.224	7.588	7.418	7.382	8.62
(S.D.)		.444	.60	.644	.788	.452	.592	.109
Other Subjects	119	5.133	4.864	6.067	5.818	4.951	5.892	7.888
(S.D.)		1.592	1.472	1.328	1.696	1.472	1.408	
Diff. in Means		2.065	2.896	1.157	1.77	2.467	1.490	.732
P.E. of Diff.		.107	.114	.112	.140	.104	.112	.085
$\frac{\text{Diff.}}{\text{P.E. of Diff.}}$		19.3	25.4	10.33	12.64	23.72	13.3	8.8

towards mathematics, but are liberal in their attitudes towards the values of the subject. The difference between the average score on Values made by the majors and the non-majors is, however, more than ten times the probable error of the difference. The difference between the mean scores made on the Purdue scale by the two groups is exceedingly small though reliable. According to this criterion of group comparison all of the scales are valid.

Another measure of the validity of the scales is the comparison of the ratings of the statements included in the scale by the two mathematics teachers with the median values assigned by the 124 judges. The total correlations showing this comparison are given in Table 3.[3] The average judgments of the two specialists

[3] These coefficients were computed from the assigned scale values and not from individual scores obtained from the scales when the statements have been assigned values by the two methods of judging. This comparison will be made later.

in the field agree very closely with the average scale values assigned by 124 judges who have only limited training in mathematics. The agreement is not so close for the statements included in the Values scale as for those included in the Interest scale. This scale is neither as reliable nor as valid as the Interest scale. The judgment of teacher B varies more from that of the group than that of teacher A. Although the comparison made here is a rough empirical one, it suggests that further investigation of the relation between arbitrarily assigned scale values and group judgments of their value should be made.

TABLE 3

CORRELATION COEFFICIENTS BETWEEN THE SCALE VALUES ASSIGNED TO THE STATEMENTS BY TWO MATHEMATICS TEACHERS AND THE MEDIAN VALUES ASSIGNED BY 124 JUDGES

Teachers	Scales	r	P.E.$_r$
A	Interest—Forms A and B	.893	.024
A	Values —Forms A and B	.911	.020
B	Interest—Forms A and B	.906	.020
B	Values —Forms A and B	.833	.035
A and B	Interest—Forms A and B	.953	.012
A and B	Values —Forms A and B	.916	.016

ATTITUDES OF INTEREST AND VALUE

Allport's conception of interests and of values, or attitudes of appreciation, suggests that there may be two related common attitude variables towards mathematics which differ in different individuals and in different groups. Some individual scores made on the Interest and the Values scales show significant differences between the attitude of interest and the attitude of appreciation. For example, one student majoring in Art made a score of 3.7 on the Interest scale and a score of 7.5 on the Values scale; one majoring in Home Economics made a score of 5.9 on the Interest scale, of 7.4 on the Values scale; and another, majoring in English made a score of 4.5 on the Interest scale and 7.0 on the Values scale. There were not enough majors in different fields to make an analysis of group trends, but the few scores chosen at random show individual differences in interest in and appreciation of the values of mathematics.

The mean differences between scores made on the Interest scales and those made on the Values scales are given in Table 4. Those not majoring in mathematics are 'neutral' both in their attitudes of interest in and appreciation of mathematics, but the score made by this group on the Values scale is higher than that made on the Interest scale and the difference is reliable. The mathematics majors, on the other hand made a slightly higher score on interest than on values, but this difference is not reliable. The scores made on both tests by this group show highly favorable attitudes towards mathematics. When the

TABLE 4

COMPARISON OF MEAN SCORES MADE ON THE INTEREST TEST AND THE VALUES TEST BY NON-MATH. AND BY MATH. MAJORS

Tests	Majoring in Other Subjects	Math. Majors	Both Groups
Values (A and B)	5.892	7.382	6.06
S.D.	1.408	.592	1.376
Interest (A and B)	4.951	7.418	5.272
S.D.	1.472	.452	1.660
Diff.	.941	.036	.788
P.E. of Diff.	.079	.088	.070
$\frac{\text{Diff.}}{\text{P.E. of Diff.}}$	11.91	.409	11.25

scores made by both groups are combined, the average score for 'values' is higher than that for 'interest.' These reliable differences between the group scores made on the two tests indicate that two attitude variables are being measured.

The correlations between the scores made on the attitude scales are given in Table 5. The coefficient between the scores made on the Interest scale (Forms A and B) and those made on the Values scale (Forms A and B) is $.609 \pm .033$, with the scores of the mathematics majors eliminated, and this coefficient becomes .717 when corrected for attenuation. When the mathematics majors are included, the total correlation is $.692 \pm .033$, which when corrected for attenuation becomes .815. The mathematics majors, with but one or two exceptions, made very high scores on both the interest and the values scales and the inclusion of their scores would necessarily raise the correlation. With

this group included, however, the evidence seems to indicate that the scales are not measuring the same attitude variables, though, as would be expected, there seems to be a good deal of overlapping between them. The total correlations between the interest and the values scales and the Purdue scale are slightly higher than forty and indicate that both scales have something in common with those measured by the Purdue scale. The higher coefficient between the Values scale and the Purdue scale is probably due to the greater number of statements of opinion about the values of mathematics included in the Purdue scale.

TABLE 5

CORRELATIONS BETWEEN SCORES MADE ON THE ATTITUDE TESTS

N	Tests	*r*	P.E.	Corrected for Attenuation
119	Interests A and B and Values A and B (Without Math. Majors)	.609	.039	.717
159	Interests A and B and Values A and B (With Math. Majors)	.692	.033	.815
119	Interests A and B and Purdue A (Without Math. Majors)	.434	.049	
159	Interests A and B and Purdue A (Both Groups)	.450	.049	
119	Values A and B and Purdue A (Without Math. Majors)	.528	.045	
159	Values A and B and Purdue A (Both Groups)	.537	.039	

All of the subjects at G. S. C. W. who took the attitude scales were women, most of whom had taught and were more mature than the average group of college students. The scores made on the attitude tests by this group of students indicate that the popular view that women in general dislike mathematics and are prejudiced against the values of the subject is erroneous. There were so few men included in the Peabody group of students that it was impossible to make any comparison of sex differences in attitudes towards mathematics.

The traditional view of mathematics is that the subject is a natural science, that it is indeed the basic tool of the natural sciences. When taught according to this conception, mathematics

is usually very abstract and is thought of by the individual as merely incidental, although necessary for the realization of more important purposes. One may not agree with Judd (26) in the conception that measurement and number relations constitute a social institution, but an analysis of mathematical processes in relation to organized group life such as he has given leads one inevitably to the conclusion that mathematics is also a social science. If conceived of in this way, the study of mathematics, as Dewey has pointed out, should result in the development of attitudes of appreciation. In a scientific age such appreciations may prove to be almost if not equally as important to the average individual as the esthetic attitudes resulting from experiences with the processes and products of the fine arts.

The relation of the attitude of appreciation or value towards mathematics to interest in the subject, to intelligence or other related attitudes or abilities will have to be determined by more accurate statistical procedures than those used in the present investigation. Thorndike (43) has recently obtained results which show that to find that one's evaluation of esthetic stimulus material agreed with the opinion of experts was, for the subjects used in his experiments, a factor in the development of attitudes of interest in or liking for stimulus material of neutral value when it was used along with the esthetic material.[4] Thurstone (49), on the basis of results obtained by Mrs. Thurstone, has shown that it is possible to apply the factor method of analysis to the problem of determining the common factors involved in a group of related attitudes. More representative data and more refined techniques of analysis may show that there is only one common factor in the attitudes measured by the Interest and the Values scales used in the present study; they may show that interest in any school subject is determined by the general patterns of attitudes or values which the individual may have

[4] The writer has not reviewed Thorndike's important experiments on the development of, and change in, wants, interests and attitudes because her own investigation gave no data on which to base an evaluation of the concept of rewards and punishment as the determining factor in the genesis and change of attitudes.

developed, such as the theoretical, the economic, the esthetic, the social, the political and the religious attitudes measured by the Allport–Vernon Test of Values (3, 9). The present study is exploratory, but the results obtained indicate the need for further investigation of the problems considered.

Summary and Conclusions

1. Women do not as a group dislike mathematics nor are they prejudiced against the values of the subject; some who major in other fields have very favorable attitudes of interest in and appreciation for the subject; some who have only a neutral or unfavorable attitude have very unfavorable attitudes both of interest and appreciation. All of the men and the women majoring in the field of mathematics had very favorable attitudes of interest and appreciation for the subject. Their attitudes of appreciation varied more than their attitudes of interest, but this result is probably due to the fact that the Values scale is not so satisfactory as the Interest scale.

2. The Interest and the Values scales are both reliable and valid, in terms of the criteria used. The results secured seem to indicate that the scales would have more value for diagnostic and research purposes than the generalized Purdue scale.

3. Though the scales developed are reliable and valid, the differences between the scores made on interest and on appreciation might disappear with the use of more analytical techniques, and so at present it is only tentatively concluded that the results seem to confirm Allport's conception that values are more intellectual and less personal to the individual than interests.

4. If the findings of the present study are confirmed by later research and it is shown that teaching mathematics in terms of the social relations represented by or made possible by mathematical processes results in the development of attitudes of appreciation or value which enable the individual better to understand and to adjust himself to his social environment, then a requirement of some study of mathematics in the general education of the average citizen will be justified on an objective basis.

ERRATUM

The first sentence of paragraph one of the Summary and Conclusions, page 176, should read as follows:

"Women do not as a group dislike mathematics nor are they prejudiced against the values of the subject; some who major in other fields have very favorable attitudes of interest in and appreciation for the subject; some who have only a neutral or unfavorable attitude of interest toward the subject have a high appreciation of its values and some have very unfavorable attitudes both of interest and appreciation."

12. Cherrington, M., and Miller, L. W. Changes in attitudes as a result of lecture and reading similar materials. *J. Soc. Psychol.*, 1933, **4**, 479–484.
13. Corey, Stephen M. Professed attitudes and actual behavior. *J. Educ. Psychol.*, 1937, **28**, 271–280.
14. Dewey, John. *Democracy and Education.* New York: Macmillan, 1916. Pp. 434.
15. Droba, D. D. Methods used for measuring public opinion. *Amer. J. Sociol.*, 1931, **37**, 410–423.
16. Droba, D. D. Education and Negro attitudes. *Sociol. Soc. Res.*, 1932, **17**, 137–141.
17. Droba, D. D. The nature of attitudes. *J. Soc. Psychol.*, 1933, **4**, 444–463.
18. Droba, D. D. Social attitudes. *Amer. J. Sociol.*, 1933, **39**, 513–524.
19. Droba, D. D. Political parties and war attitudes. *J. Abn. and Soc. Psychol.*, 1934, **28**, 468–472.
20. Ferguson, L. W. The influence of individual attitudes on construction of an attitude scale. *J. Soc. Psychol.*, 1935, **6**, 115–117.
21. Fryer, Douglas. *The Measurement of Interests.* New York: Henry Holt, 1931. Pp. xxxvi+488.
22. Gardner, Iva Cox. The effect of a group of social stimuli upon attitudes. *J. Educ. Psychol.*, 1935, **26**, 471–478.
23. Guilford, J. P. *Psychometric Methods.* New York: McGraw-Hill Book Company, Inc., 1936. Chap. 5.
24. Hinckley, E. D. The influence of individual opinion on the construction of an attitude scale. *J. Soc. Psychol.*, 1932, **3**, 283–295.
25. Holzinger, Karl J. *Statistical Methods in Education.* Boston: Ginn and Company, 1928. Pp. viii+372.
26. Judd, Charles H. *The Psychology of Social Institutions.* New York: Macmillan, 1931. Chaps. 5, 6, and 7.

2. The Interest and the Values scales are both reliable and valid, in terms of the criteria used. The results secured seem to indicate that the scales would have more value for diagnostic and research purposes than the generalized Purdue scale.

3. Though the scales developed are reliable and valid, the differences between the scores made on interest and on appreciation might disappear with the use of more analytical techniques, and so at present it is only tentatively concluded that the results seem to confirm Allport's conception that values are more intellectual and less personal to the individual than interests.

4. If the findings of the present study are confirmed by later research and it is shown that teaching mathematics in terms of the social relations represented by or made possible by mathematical processes results in the development of attitudes of appreciation or value which enable the individual better to understand and to adjust himself to his social environment, then a requirement of some study of mathematics in the general education of the average citizen will be justified on an objective basis.

Bibliography

1. ACHESON, EUNICE M. A study of graduate women's reactions and opinions on some modern social attitudes and practices. *J. Abn. and Soc. Psychol.*, 1933, **28**, 42–63.
2. ALLPORT, F. H. Teleonomic description in the study of personality. *Char. and Pers.*, 1937, **5**, 202–214.
3. ALLPORT, G. W. Attitudes. Chap. 7 in *Handbook of Social Psychology* by Carl Murchison. Worcester: Clark University Press, 1935.
4. BAIN, R. An attitude on attitude research. *Amer. J. Sociol.*, 1927, 33, 940–957.
5. BAIN, R. Theory and measurement of attitude and opinions. *Psychol. Bull.*, 1930, **27**, 357–379.
6. BOLTON, EURI BELLE. Measuring specific attitudes towards the social rights of the Negro. *J. Abn. and Soc. Psychol.*, 1937, **31**, 384–397.
7. BOLTON, EURI BELLE. Effect of knowledge upon attitudes towards the Negro. *J. Soc. Psychol.*, 1935, **6**, 68–90.
8. CAJORI, FLORIAN. *Mathematics in a Liberal Education.* Boston: Christopher Publishing House, 1928. Pp. 169.
9. CANTRIL, H. General and specific attitudes. *Psychol. Monog.*, 1932, **42**, No. 192. Pp. 109.
10. CARLSON, HILDING B. Attitudes of undergraduate students. *J. Soc. Psychol.*, 1934, **5**, 202–213.
11. CASWELL, HOLLIS L., and CAMPBELL, DOAK S. *Curriculum Development.* New York: American Book Company, 1935. Pp. xvii+600.
12. CHERRINGTON, M., and MILLER, L. W. Changes in attitudes as a result of lecture and reading similar materials. *J. Soc. Psychol.*, 1933, **4**, 479–484.
13. COREY, STEPHEN M. Professed attitudes and actual behavior. *J. Educ. Psychol.*, 1937, **28**, 271–280.
14. DEWEY, JOHN. *Democracy and Education.* New York: Macmillan, 1916. Pp. 434.
15. DROBA, D. D. Methods used for measuring public opinion. *Amer. J. Sociol.*, 1931, **37**, 410–423.
16. DROBA, D. D. Education and Negro attitudes. *Sociol. Soc. Res.*, 1932, 17, 137–141.
17. DROBA, D. D. The nature of attitudes. *J. Soc. Psychol.*, 1933, **4**, 444–463.
18. DROBA, D. D. Social attitudes. *Amer. J. Sociol.*, 1933, **39**, 513–524.
19. DROBA, D. D. Political parties and war attitudes. *J. Abn. and Soc. Psychol.*, 1934, **28**, 468–472.
20. FERGUSON, L. W. The influence of individual attitudes on construction of an attitude scale. *J. Soc. Psychol.*, 1935, **6**, 115–117.
21. FRYER, DOUGLAS. *The Measurement of Interests.* New York: Henry Holt, 1931. Pp. xxxvi+488.
22. GARDNER, IVA COX. The effect of a group of social stimuli upon attitudes. *J. Educ. Psychol.*, 1935, **26**, 471–478.
23. GUILFORD, J. P. *Psychometric Methods.* New York: McGraw-Hill Book Company, Inc., 1936. Chap. 5.
24. HINCKLEY, E. D. The influence of individual opinion on the construction of an attitude scale. *J. Soc. Psychol.*, 1932, **3**, 283–295.
25. HOLZINGER, KARL J. *Statistical Methods in Education.* Boston: Ginn and Company, 1928. Pp. viii+372.
26. JUDD, CHARLES H. *The Psychology of Social Institutions.* New York: Macmillan, 1931. Chaps. 5, 6, and 7.

27. KATZ, D., and ALLPORT, F. H. *Student Attitudes.* Syracuse: Craftsman Press, 1931. Pp. iii+408.
28. KATZ, D. Attitude measurement as a method in social psychology. *Social Forces,* 1937, **15**, 479–482.
29. KNOWER, F. H. Experimental studies of changes in attitudes. I. A study of the effect of oral argument on changes of attitudes. *J. Soc. Psychol.,* 1935, **6**, 315–345.
30. LASHLEY, K. D. Basic neural mechanisms in behavior. *Psychol. Rev.,* 1930, **37**, 1–24.
31. LIKERT, R. A. A technique for the measurement of attitude. *Arch. Psychol.,* 1932, **22**, 5–55.
32. MOORE, GWYN, and GARRISON, K. C. A comparative study of social and political attitudes of college students. *J. Abn. and Soc. Psychol.,* 1932, **27**, 195–208.
33. MURPHY, GARDNER, and MURPHY, LOIS BARCLAY. *Experimental Social Psychology.* New York: Harper & Bros., 1931. Chap. 2.
34. PETERSON, J. Aspects of learning. *Psychol. Rev.,* 1935, **42**, 1–27.
35. PINTNER, R., and FORLANO, G. The influence of attitude upon scaling attitude items. *J. Soc. Psychol.,* 1937, **8**, 39–45.
36. REMMERS, H. H. Studies in attitudes. *Bull. Purdue Univ. Stud. Higher Educ.,* 26, 1934. Pp. 112.
37. REMMERS, H. H. Further studies in attitudes. Series II. *Bull. Purdue Univ. Stud. Higher Educ.,* 31, 1934. Pp. 298.
38. ROSANDER, A. C. An attitude scale based upon behavior situations. *J. Soc. Psychol.,* 1937, **8**, 3–15.
39. RUNDQUIST, EDWARD A., and SLETTO, RAYMOND F. *Personality in the Depression.* Minneapolis: University of Minn. Press, 1936. Pp. xxii+398.
40. SEASHORE, R. H., and HEVNER, KATE. A time-saving device for the construction of attitude scales. *J. Soc. Psychol.,* 1933, **4**, 366–372.
41. SOWARDS, GENEVIEVE S. A study of the attitudes of college students. *J. Abn. and Soc. Psychol.,* 1934, **29**, 328–340.
42. TELFORD, C. W. An experimental study of some factors influencing the attitudes of college students. *J. Soc. Psychol.,* 1934, **5**, 421–428.
43. THORNDIKE, E. L. *The Psychology of Wants, Interests and Attitudes.* New York: D. Appleton-Century, 1935. Pp. x+301.
44. THURSTONE, L. L., and CHAVE, E. J. *The Measurement of Attitude.* Chicago: Univ. Chicago Press, 1929. Pp. 96.
45. THURSTONE, L. L. An experimental study of national preferences. *J. Gen. Psychol.,* 1928, **1**, 405–425.
46. THURSTONE, L. L. Attitudes can be measured. *Amer. J. Sociol.,* 1928, **33**, 529–554.
47. THURSTONE, L. L. The measurement of change in social attitudes. *J. Soc. Psychol.,* 1931, **2**, 230–235.
48. THURSTONE, L. L. The measurement of social attitudes. *J. Abnor. and Soc. Psychol.,* 1931, **26**, 249–269.
49. THURSTONE, L. L. The vectors of mind. *Psychol. Rev.,* 1934, **41**, 1–32.
50. WILLIAMSON, E. G., and DARLEY, J. G. The measurement of social attitudes of college students: I. standardization of test and results of a survey. *J. Soc. Psychol.,* 1937, **8**, 219–229.
51. WILLIAMSON, E. G., and DARLEY, J. G. The measurement of social attitudes of college students: II. validation of two attitude tests. *J. Soc. Psychol.,* 1937, **8**, 231–242.

APPENDIX I

Q-VALUES (*i.e.*, $\frac{Q_3-Q_1}{2}$) FOR THE EIGHTY-TWO STATEMENTS INCLUDED IN THE FOUR FORMS OF THE SCALES

Interest A	Interest B	Values A	Values B
.5	1.0	.85	.95
.7	.4	.35	.4
1.35	1.15	1.0	1.15
1.2	.35	.9	1.2
.3	.55	.65	1.05
.7	1.05	.9	.45
1.5	1.7	1.55	1.55
1.1	1.0	1.2	1.0
1.2	.75	.55	.4
1.0	1.70	1.15	1.35
1.0	.4	.75	.65
.3	.8	.80	.95
.85	1.0	1.3	.95
1.35	1.15	.48	1.55
1.05	.6	1.2	1.05
1.2	1.1	1.1	1.3
1.35	1.5	1.1	1.3
1.6	1.2	.55	1.35
1.05	1.1	1.7	.48
1.00	1.1	1.1	1.45
1.1	1.15		

Appendix II

1. FORM A OF THE 'INTEREST' SCALE

Name: *Date:*
Class: *Major:*

Attitudes Toward Mathematics—Form A

Put a check mark (✓) if you agree with the statement.
Put a cross (×) if you disagree with the statement.

( ) 1. My attitude toward mathematics is very unfavorable and that hinders my progress in the study.
( ) 2. Every phase of mathematics appeals to me in every way.
( ) 3. Mathematics is not so difficult to me as other school subjects.
( ) 4. One who has confidence in his ability to think should have no fear of mathematics.
( ) 5. Mathematics is about equal in difficulty to other school subjects.
( ) 6. I like mathematics very much.
( ) 7. I am indifferent to mathematics.
( ) 8. I always approach the study of any subject in mathematics with a tense, unpleasant feeling.
( ) 9. I have had many unpleasant experiences in connection with the study of mathematics.
( ) 10. I am especially interested in algebra, arithmetic, and geometry.
( ) 11. I am interested only in practical mathematics.
( ) 12. I neither like nor dislike mathematics.
( ) 13. Nothing thrills me more than to solve a really difficult problem in mathematics.
( ) 14. I like mathematics fairly well, but it becomes very laborious at times.
( ) 15. I like to work with anything that is in any way connected with figures.
( ) 16. Mathematics is difficult for me because I memorized it instead of learning the principles.
( ) 17. The difficulty in mathematics lies in its abstraction.
( ) 18. I think that students who dislike mathematics should not be required to take it.
( ) 19. Mathematics is hard but interesting.
( ) 20. Mathematics requires entirely too much time for most students.
( ) 21. Mathematics is hard for me because of my lack of interest in it.

2. FORM B OF THE 'INTEREST' SCALE

Name: *Date:*
Class: *Major:*

Attitudes Toward Mathematics—Form B

Put a check mark (✓) if you agree with the statement.
Put a cross (×) if you disagree with the statement.

( ) 1. Mathematics is very monotonous.
( ) 2. I love mathematics and think it the most interesting subject I have ever studied.

( ) 3. I enjoy algebra and geometry more than any other mathematics.
( ) 4. I neither like nor especially dislike mathematics.
( ) 5. Mathematics is my favorite study.
( ) 6. Mathematics is more difficult than language and literature and less difficult than science.
( ) 7. My dislike for mathematics was caused by a constant fear of failing the subject.
( ) 8. Mathematics involves too much technical work to be interesting.
( ) 9. It gives me great pleasure to work with fractions and equations in mathematics.
( ) 10. Mathematics requires a great deal of reasoning and work.
( ) 11. Sometimes mathematics interests me; at other times it doesn't.
( ) 12. I like mathematics and find it very interesting.
( ) 13. I am interested in mathematics, but I would not care to major in this field.
( ) 14. I do not dislike arithmetic and algebra, but I dislike geometry.
( ) 15. Mathematical problems are fascinating.
( ) 16. I dislike geometry because I did not have a good foundation for it.
( ) 17. I like arithmetic, but I do not like the other courses in mathematics.
( ) 18. I had rather read a book than solve a problem in mathematics.
( ) 19. It is very trying on the nerves to study mathematical problems for any length of time.
( ) 20. Mathematics appeals to the imagination.
( ) 21. My interest in mathematics is thin.

3. FORM A OF THE 'VALUES' SCALE

Name: *Date:*
Class: *Major:*

Attitudes Toward Mathematics—Form A

Put a check mark (√) if you agree with the statement.
Put a cross (×) if you disagree with the statement.

( ) 1. Mathematics is just as important to one as English.
( ) 2. Mathematics is the most important subject because it makes possible an understanding of the natural laws of the universe.
( ) 3. Few girls should be required to labor with algebra.
( ) 4. Required mathematics is a waste of time.
( ) 5. Mathematics is necessary in all phases of life.
( ) 6. The value of mathematics is very limited.
( ) 7. Only practical everyday problems should be included in the study of mathematics.
( ) 8. Complex mathematical formulae have no value whatsoever.
( ) 9. Some forms of mathematics are valuable, but some are not.
( ) 10. Number relationships are so likely to be misinterpreted that mathematics has little value for the social sciences.
( ) 11. Mathematics is very valuable because we use it in everyday life.
( ) 12. Mathematics should have a more important place in the curriculum than it has at present.
( ) 13. Girls should not be required to take mathematics.
( ) 14. The specialized fields of mathematics have no value to the average person.

( ) 15. Simple mathematics is very valuable.
( ) 16. Everyone should have some general knowledge of mathematics.
( ) 17. The value of mathematics varies according to the vocation of the person concerned.
( ) 18. Mathematics has about the same value as other school subjects.
( ) 19. Mathematics is as valuable as most other subjects.
( ) 20. Every student should be required to take a course in the mathematics of finance.

4. FORM B OF THE 'VALUES' SCALE

Name: *Date:*
Class: *Major:*

Attitudes Toward Mathematics—Form B

Put a check mark (✓) if you agree with the statement.
Put a cross (×) if you disagree with the statement.

( ) 1. Mathematics is needed for almost every other course one studies.
( ) 2. The value of mathematics is unlimited because it is essential for so many fields today.
( ) 3. Only a limited amount of mathematics should be taught.
( ) 4. Algebra is a useless waste of time for the majority of boys and for most girls.
( ) 5. A knowledge of mathematics is basic to all other knowledge.
( ) 6. Mathematics has very little value in comparison with other school subjects.
( ) 7. Higher courses in mathematics are needed by only a very few people.
( ) 8. The average person needs no higher mathematics than seventh grade arithmetic; higher mathematics is useless.
( ) 9. Mathematics has more value than some subjects, but less than others.
( ) 10. Mathematics makes progress in all sciences possible.
( ) 11. Mathematics does not help one learn to deal with people.
( ) 12. Mathematics is used more than any subject except reading and writing.
( ) 13. The complicated and impractical courses in mathematics have no value whatsoever.
( ) 14. Algebra and geometry should be put strictly on the elective basis both in high school and in college.
( ) 15. Mathematics is more valuable for high school and college students than foreign languages.
( ) 16. Business arithmetic has more value than any subject in the high school curriculum.
( ) 17. Mathematics is used more in business than any other subject.
( ) 18. Arithmetic and geometry have practical value, but algebra has none.
( ) 19. Some forms of mathematics are valuable, but others are not valuable at all.
( ) 20. Certain courses in mathematics are valuable.

A STUDY OF THE INTERESTS OF COLLEGE STUDENTS

K. C. GARRISON

STATE COLLEGE, UNIVERSITY OF NORTH CAROLINA

Scientific studies relating to a more accurate procedure for the measurement of interests have been under way in this country for somewhat over a decade. Psychometric instruments designed to measure qualities other than education and intelligence have grown to such a degree of importance that many colleges, universities, business and industrial establishments use them in conducting more efficient personnel work with the students or workers involved. Although vocational and occupational problems have stimulated a large amount of the work that has been done in the measurement of interests, Fryer emphasizes the importance of the cultural objective for the measurement of interests. He says:

> "To view interest measurement in true perspective we must start with an assumption different from the one current in this commercial age in which successful accomplishment is the criterion of measurement. Happy accomplishment is the foundation of a modern individualistic philosophy. Interest measurement is concerned with a distribution of interests which this philosophy assumes as the basis of happiness. The significance of interest measurement lies not in its relation to social efficiency, but rather in its measurement of a cultural development which is related to social happiness (**8**, p. v)."

This investigation presents a further effort to answer certain questions relative to the *interests* of college students. The following are some of the major problems with which this study is concerned: (1) What differences in interest exist among students registering in different divisions of work at a State land-grant college? (2) What is the nature of any difference existing between the interests of freshmen and upper-classmen? (3) How do the interests of students with an agricultural background compare with those from an engineering and business background? (4) What uses can be made of the results of a test thus constructed in aiding in personnel work with students?

Historical Review

1. *Methods of Measuring Interests.* The chief methods that have been used in the measurement of interests are: (1) inventories or questionnaires, (2) rating scales, (3) information tests, and (4) free association tests. Fryer classifies the first two of these as subjective procedures, since they represent an estimate made by the individual of his own reaction tendencies (7). The third and fourth methods are considered objective procedures since they are not so dependent upon subjective estimates by the self or by an observer. However, within recent years a number of students of education and psychology have concerned themselves with problems pertaining to increasing the validity and reliability of questionnaire and rating scale procedures. In the case of the rating scale Remmers (18) and Adams (1) suggested inserting judgments of objective qualities when possible as a means of eliminating the halo effect and to give a more exact measure for comparative purposes. Thurstone's attempts (25) to find the "primary factors" which make up the pattern of one's interests represents an attempt to simplify and make more efficient the questionnaire and interview procedures for evaluating interests.

One of the earliest scientific approaches to the problem of analyzing one's own likes and dislikes is the "Analysis of Work Interests," by Miner (15). In this he developed at Carnegie Institute of Technology in 1918 a questionnaire which he has revised from time to time. This is particularly suitable to high school and junior college students and serves to focus attention on significant aspects of occupational interests. Strong's *Vocational Interest Blank* (19) is the best known standardized inventory available for ascertaining the similarity between one's interests and those of individuals actually engaged in specific occupations on the professional level. This test is an expansion and revision of a previous inventory developed by Cowdery (8). Strong's original interest blank consisted of 263 items which sampled a range of interests, such as: occupations, amusements, school subjects, activities, peculiarities of people, order of pref-

erences, and comparison of interests between items. Keys have been prepared for scoring the revised blank for twenty-four or more occupations.

The rating scale method has not been used so extensively as the questionnaire. Probably the studies by Cox best represent the value of the use that might be made of this technique (5). She used a seven-point scale ranging from plus 3 to minus 3. A rating of plus 3 was assigned as the highest degree, while a rating of minus 3 was assigned as the lowest. Intermediate ratings were thus assigned according to the intermediate degree of a particular interest present. Other investigators have used this technique in studying different aspects or types of interest, such as: intellectual interests, social interests, play interests, activity interests, and professional interests.

A number of investigators have experimented with the information test to measure interests of groups and individuals. It has been constructed with the aim of covering casual and extensive information in a specialized field as, for example, the mechanical field. Terman's information test of play interest is a multiple choice test covering various types of games (22, p. 648). These tests are recognized as more objective in nature than the inventories and rating scales, and they yield a higher degree of reliability upon statistical treatment.

The principle of the free association test is to measure interests by the type of responses given to specific key words. These tests are based upon the assumption that such responses will reveal the nature of the interests present. In the investigation by Jennie B. Wyman (28) this method was employed. She devised an association test consisting of 120 stimulus words to be presented verbally and with responses to be given in writing. This test was designed for young people between the ages of eight and fifteen, and covers interest reactions in three fields: intellectual, social, and activity.

The present status of interest measurements is only a little more advanced than was that for the measurement of intelligence during the early part of the present century. The methods referred to on the previous pages show the manner in which this

problem is being attacked. What appears to be most needed is an analysis, a modification, and a refinement of the procedures that have thus far been developed. The more subjective nature of interests presents a problem in its measurement that is not quite so much in evidence in the case of the measurement of abilities. Fryer (7, 8) suggested this a few years ago when he wrote:

> " Abilities have always been predominantly objective in their definition, even prior to their measurement, while the early conceptions of interests are all subjective. In all the fields of interests there is the additional problem of finding out exactly what are objective interests, or what is the objective interest, or what are known as subjective interests, so that the measurement of interests may become an objective problem."

The objective measurement of interests is still to a large degree in the experimental stage. It is in the measurement of subjective interests that well defined techniques have been developed. Some of these techniques are especially well defined and quite complicated in nature. These subjective procedures give promise of contributing something of value in the measurement of human traits. Until valid objective procedures are developed and refined, these procedures will continue to offer a fundamental contribution to measurement.

2. *Results of Typical Earlier Studies.* In order to present a better basis for organizing the materials for this study, it is worthwhile to review some of the major studies of recent date that relate to the problems with which this investigation is concerned. In 1931 Fryer (8) presented a rather thorough review of interest studies and the following year gave some very constructive procedures for validating measures of interest (9).

Remmers (17) made a comparative study of the interests of students of Engineering and Agriculture. A discriminating diagnostic test for the measurement of interests was developed by means of tabulating those items which showed a statistically significant difference between these two groups for " like " or " dislike." This investigation was based on tests administered to 100 students of Agriculture and to 200 students of Engineering. The responses were then tabulated and studied for their differentiating values. From this study he concludes: " It is possible to measure with a fair degree of reliability the diver-

gence of interests as between students who enroll in engineering and those who enroll in agriculture."

Bervard O. Nemoitin (16) investigated the relation between interest and achievement. The data of his study were gathered by means of a questionnaire and the use of school records of the students. He found that the degrees of relationship between ability in high school courses "liked best," "liked second best," "disliked most," "disliked next as much," and average ability for high school courses are expressed by the correlation coefficients .60 ± .04, .49 ± .04, .58 ± .04, and .57 ± .04, when the data obtained from 150 high school seniors are considered. The relationship between interest and ability was found to become more variable and hence less reliable as the degree of interest considered moved from the extremes.

One hundred and fifty-six seniors of the class of 1927 filled out Strong's *Vocational Interest Blank* a few weeks before graduation (20). In January 1928 they reported their occupational choices. Comparison of these two sets of data shows:

> "46% have entered the occupation on which they scored highest in the test."
> "20% have entered the occupation on which they scored second highest."
> "11% entered the occupation on which they scored third highest."
> "Therefore, 77% have entered the occupation on which they scored first, second, or third highest from among fifteen different occupations. Only 18% have entered an occupation for which, according to the test, they have no interest. When several modifying factors are taken into consideration, as outlined, the results appear still better. Considering the fact that many of these men will change their occupation in the next few years and that some have entered occupations they are not interested in due to family and financial considerations, these results are about as satisfactory as can be expected from a valid vocational guidance test."

The study by Goodfellow (11) is of interest in this connection. The Strong Vocational Interest Blank was used with prospective teachers and comparisons were made between the A (possessing interests of successful teachers) and the C (not possessing such interests) groups. The A group was superior in average grades earned. In the case of the 18 women given the test, they were less introverted and more stable emotionally, while no reliable personality differences were noted for the 12 men taking the test.

Strong (21), in his study of the maturity of interests found

that there is a decided change from the age interval of 15 to 22. In the ten year intervals from 25 to 55 the changes were not pronounced. Lehman and Witty (14) found that permanency of interest does not exist in many of the vocational preferences of boys. They concluded that perhaps no phase of human nature is subject to such marked changes as that reflected in the vocational interests and preferences of growing boys. Thorndike (23) concluded in an early study of the permanence of interests that it would be hard to discover "any feature of a human being which was a much more permanent fact of his nature than his relative

TABLE 1

NUMBER AND PER CENT OF STUDENTS ENTERING THE OCCUPATION IN WHICH THEY ARE RATED FIRST, SECOND, THIRD, FOURTH, OR FIFTH, ACCORDING TO THE INTEREST TEST *(After Strong)*

Entering the occupation in which they receive:	Entire 156 Students		41 Students Not Sure of Their Choice		115 Students Sure of Their Choice	
	No.	%	No.	%	No.	%
The highest rating	71	46	14	34	57	50
The 2nd highest rating	31	20	7	17	24	21
The 3rd highest rating	17	11	4	10	13	11
The 4th highest rating	7	4	2	5	5	4
The 5th highest rating	2	1	0	0	2	2
A-C Rating	28	18	14	34	14	12
Total	156	100	41	100	115	100

degree of interest in different lines of thought and action." He also stated that interest and ability are so closely related that "either may be used as a sympton of the other almost as well as for itself."

Florence M. Young (29) investigated the causes for loss of interest in high school subjects. She obtained her results from a questionnaire which was given to 651 college women attending State Teachers College, Athens, Georgia. A number of the outstanding reasons were as follows: failure to see a need for the subject (29%); uninteresting material (24%); monotonous methods (23%); inability of the teacher to "put it across" (23%); lack of foundation (22%); difficulty of material (20%); incompetent or ignorant teacher (17%).

Tests and Subjects Used

The test used for measuring interests in this study is an outgrowth of previous studies begun in the fall of 1935. The original test employed the procedure of paired comparisons which was used in one section of Strong's *Vocational Interest Blank*. This paired comparisons technique was used by Furfey (10) in his scale for measuring developmental age in boys. The original inventory consisted of the four sections which are presented in the revised and enlarged form; however, in the original there are only seventy-five comparisons presented, while in the revised form there are ninety-five. As a result of the trial experiences, many changes were made in the items presented in the paired comparisons. The items of the revised and enlarged test which was used for gathering data for this study are listed in Table 2. The directions for performing the various parts of the test are quite similar in nature. In the case of A "Occupations," the following directions were presented along with the test:

> "In each pair choose the occupation you think you would like best. Consider only whether or not you like the type of work; do not consider the money earned by persons in each occupation or the social standing they usually have. Would you like the kind of work involved in the occupation if everything else were satisfactory?"

Data were gathered on 320 students at North Carolina State College. Complete data were available on 166 upper-classmen distributed as follows: 42 engineering, 46 textiles, 29 business, and 20 vocational education. Previous investigations of tests of this type have shown them to have a high degree of reliability. Upon giving this test a second time, approximately five months after the first one, to eleven students, it was found that the average number of changes was a little less than two.

Results

Lacking a standard for use in making comparisons of interests, the principle of relativity was used. Comparisons were made between the interests of upper-classmen registered in the different divisions of the college. These comparisons are based upon the per cent of students choosing the one item in preference to

TABLE 2

CHOICES MOST FREQUENTLY MADE BY STUDENTS FROM DIFFERENT DIVISIONS OF NORTH CAROLINA STATE COLLEGE

		Engineering	Textiles	Forestry	Agriculture	Voc. Educa.	Business	Freshmen	Upperclassmen
A. *Occupations:*									
1. An express clerk	or, 2. An auto mechanic	(2)	(2)	(2)	(1)	(1)	(1)	(2)	(–)
1. A rancher	or, 2. A truck driver	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. A mail carrier	or, 2. A dairyman	(1)	(1)	(2)	(2)	(1)	(1)	(–)	(1)
1. A telegraph operator	or, 2. A bank teller	(–)	(2)	(–)	(2)	(2)	(2)	(1)	(2)
1. An electrician	or, 2. A forest ranger	(1)	(2)	(2)	(2)	(2)	(1)	(2)	(–)
1. An author	or, 2. A barber	(2)	(2)	(1)	(1)	(2)	(2)	(2)	(2)
1. A chemist	or, 2. A civil engineer	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. A stock broker	or, 2. A doctor	(–)	(1)	(2)	(2)	(2)	(1)	(2)	(–)
1. A mechanic	or, 2. A cowboy	(1)	(1)	(2)	(2)	(1)	(1)	(1)	(1)
1. A forest ranger	or, 2. A bank teller	(1)	(–)	(1)	(1)	(1)	(2)	(1)	(–)
1. A railroad engineer	or, 2. A college professor	(1)	(1)	(1)	(2)	(2)	(2)	(1)	(1)
1. A farmer	or, 2. A merchant	(2)	(2)	(1)	(1)	(2)	(2)	(1)	(–)
1. A librarian	or, 2. An auto salesman	(2)	(2)	(2)	(2)	(–)	(2)	(2)	(2)
1. A scientist	or, 2. A civil engineer	(2)	(2)	(2)	(2)	(2)	(2)	(–)	(2)
1. A salesman	or, 2. A forest ranger	(2)	(2)	(2)	(2)	(2)	(1)	(2)	(2)
1. A stockbroker	or, 2. A dairyman	(1)	(1)	(2)	(2)	(2)	(1)	(–)	(1)
1. An electrician	or, 2. A merchant	(1)	(2)	(1)	(2)	(2)	(2)	(1)	(2)
1. An author	or, 2. A civil engineer	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. A bank teller	or, 2. A farmer	(–)	(1)	(2)	(2)	(2)	(1)	(–)	(1)
1. A chemist	or, 2. A dairyman	(1)	(1)	(–)	(2)	(2)	(1)	(1)	(1)

Note: Differences were considered negligible and given a zero rating when they were less than 10% in favor of either choice.

TABLE 2—*Continued*

		Engineering	Textiles	Forestry	Agriculture	Voc. Educa.	Business	Freshmen	Upperclassmen
A. *Occupations:* (Continued)									
1. An electrician	or, 2. A salesman	(1)	(-)	(1)	(1)	(1)	(2)	(1)	(1)
1. A librarian	or, 2. An electrician	(2)	(2)	(2)	(-)	(-)	(1)	(1)	(2)
1. A telegraph operator	or, 2. A lawyer	(1)	(2)	(-)	(2)	(2)	(2)	(2)	(2)
1. A farmer	or, 2. A teacher	(1)	(2)	(1)	(1)	(2)	(2)	(-)	(-)
1. A dairyman	or, 2. An auto mechanic	(2)	(2)	(2)	(1)	(1)	(-)	(1)	(2)
1. A farm manager	or, 2. A store manager	(2)	(2)	(1)	(1)	(1)	(2)	(1)	(2)
1. A newspaper reporter	or, 2. A furniture maker	(-)	(1)	(1)	(1)	(1)	(1)	(2)	(1)
1. A match maker	or, 2. A county farm agent	(1)	(-)	(2)	(2)	(2)	(2)	(1)	(2)
1. A bookkeeper	or, 2. A high school teacher	(2)	(1)	(2)	(2)	(2)	(1)	(-)	(-)
1. A poultryman	or, 2. A stenographer	(-)	(2)	(1)	(1)	(1)	(2)	(2)	(-)
B. *School Subjects:*									
1. Science	or, 2. English	(1)	(1)	(1)	(1)	(-)	(-)	(-)	(-)
1. Art	or, 2. Manual training	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Music	or, 2. Biology, nature study	(1)	(2)	(2)	(2)	(2)	(-)	(1)	(2)
1. Language	or, 2. Penmanship	(2)	(1)	(1)	(1)	(1)	(1)	(-)	(1)
1. Algebra	or, 2. Physical education	(-)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Manual training	or, 2. History	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. Agriculture	or, 2. Geometry	(2)	(2)	(1)	(1)	(1)	(2)	(2)	(2)
1. Shop work	or, 2. Music	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. Biology, nature study	or, 2. Literature	(1)	(-)	(1)	(1)	(1)	(-)	(1)	(1)
1. Agriculture	or, 2. Shop work	(2)	(2)	(1)	(1)	(1)	(2)	(2)	(2)
1. Physical education	or, 2. Literature	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. Bookkeeping	or, 2. Nature study	(2)	(1)	(2)	(2)	(2)	(1)	(2)	(-)
1. English	or, 2. Manual training	(2)	(2)	(2)	(2)	(2)	(-)	(2)	(2)
1. Agriculture	or, 2. Art	(2)	(2)	(1)	(1)	(1)	(-)	(1)	(-)
1. Physical education	or, 2. General science	(2)	(-)	(1)	(1)	(1)	(1)	(-)	(2)

TABLE 2—*Continued*

		Engineering	Textiles	Forestry	Agriculture	Voc. Educa.	Business	Freshmen	Upperclassmen
B. *School Subjects:* (Continued)									
1. Agriculture	or, 2. Typing	(–)	(2)	(1)	(1)	(1)	(2)	(2)	(2)
1. Literature	or, 2. Shop work	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Manual training	or, 2. Science	(2)	(2)	(2)	(–)	(–)	(–)	(1)	(–)
1. Arithmetic	or, 2. English	(1)	(1)	(–)	(1)	(–)	(1)	(1)	(1)
1. Biology	or, 2. Bookkeeping	(1)	(2)	(1)	(1)	(1)	(1)	(1)	(–)
C. *Things to have:*									
1. A pet dog	or, 2. A tennis racket	(1)	(1)	(1)	(1)	(1)	(–)	(1)	(1)
1. An electric train	or, 2. A radio	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. A new encyclopaedia	or, 2. A rifle	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. A camera	or, 2. A fountain pen	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. A new suit	or, 2. A tool chest	(1)	(1)	(1)	(–)	(–)	(1)	(1)	(1)
1. A beautiful horse	or, 2. A radio set	(1)	(1)	(1)	(1)	(2)	(–)	(1)	(1)
1. An architectural project	or, 2. A new drawing set	(2)	(2)	(2)	(1)	(1)	(2).	(2)	(2)
1. A rifle	or, 2. A beautiful horse	(2)	(2)	(–)	(2)	(2)	(2)	(2)	(2)
1. A new encyclopaedia	or, 2. A pet dog	(2)	(2)	(2)	(–)	(2)	(2)	(2)	(2)
1. A tennis racket	or, 2. A tool chest	(–)	(1)	(–)	(2)	(2)	(–)	(2)	(–)
1. A new suit	or, 2. A new encyclopaedia	(1)	(1)	(1)	(1)	(2)	(1)	(1)	(1)
1. A tool chest	or, 2. An agricultural project	(1)	(1)	(2)	(2)	(2)	(1)	(1)	(1)
1. A new fountain pen	or, 2. A pet dog	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. A new school bag	or, 2. A rifle	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. A camera	or, 2. A tool chest	(–)	(1)	(1)	(–)	(–)	(1)	(1)	(1)
1. A prize winning pig	or, 2. A new typewriter	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. A dictionary	or, 2. A handsaw	(1)	(1)	(1)	(–)	(1)	(1)	(–)	(1)
1. An emery rock	or, 2. A registered calf	(1)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Some Hungarian pigeons	or, 2. A new rifle	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. An adding machine	or, 2. A set of good books	(–)	(2)	(2)	(2)	(2)	(2)	(2)	(2)

TABLE 2—*Continued*

		Engineering	Textiles	Forestry	Agriculture	Voc. Educa.	Business	Freshmen	Upperclassmen
D. *Things to do:*									
1. Repair a clock	or, 2. Make a speech	(1)	(1)	(1)	(–)	(–)	(1)	(1)	(1)
1. Handle horses	or, 2. Sell merchandise	(–)	(2)	(1)	(1)	(1)	(2)	(1)	(–)
1. Make a speech	or, 2. Work with a radio	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Go to school	or, 2. Work on a farm	(1)	(1)	(1)	(2)	(2)	(1)	(1)	(1)
1. Build a model plane	or, 2. Train horses	(1)	(1)	(2)	(2)	(2)	(1)	(2)	(–)
1. Read biographies	or, 2. Visit factories	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Adjust carburetor	or, 2. Work in a garden	(1)	(1)	(2)	(2)	(–)	(1)	(1)	(1)
1. Write a science report	or, 2. Clerk in a store	(–)	(2)	(–)	(–)	(2)	(2)	(2)	(2)
1. Work with animals	or, 2. Make a speech	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. Attend club meetings	or, 2. Ride horses	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Keep an expense account	or, 2. Write stories	(1)	(1)	(1)	(1)	(–)	(1)	(–)	(1)
1. Read literature	or, 2. Work with tools	(2)	(2)	(2)	(2)	(2)	(–)	(2)	(2)
1. Working on a farm	or, 2. Working in a factory	(2)	(2)	(1)	(1)	(1)	(–)	(–)	(2)
1. Planting a crop	or, 2. Building a radio set	(2)	(2)	(1)	(1)	(1)	(2)	(2)	(2)
1. Clerk in a store	or, 2. Study your lesson	(2)	(1)	(2)	(–)	(–)	(1)	(1)	(1)
1. Looking at shop windows	or, 2. Visit a dairy	(2)	(1)	(2)	(2)	(2)	(1)	(2)	(1)
1. Visit factories	or, 2. Climb hills	(–)	(1)	(2)	(2)	(2)	(1)	(1)	(1)
1. Collect stamps	or, 2. Work with animals	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Sell merchandise	or, 2. Read biographies	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. Climb hills	or, 2. Collect stamps	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. Write a story	or, 2. Make a bookcase	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Work arithmetic	or, 2. Ready a story	(–)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Feed chickens	or, 2. Type a letter	(–)	(2)	(1)	(1)	(1)	(2)	(2)	(2)
1. Display merchandise	or, 2. Prepare agricultural exhibit	(1)	(1)	(2)	(2)	(2)	(1)	(–)	(1)
1. Tighten bearing on a car	or, 2. Dehorn a cow	(1)	(1)	(1)	(2)	(2)	(1)	(1)	(1)

TABLE 2—*Continued*

		Engineering	Textiles	Forestry	Agriculture	Voc. Educa.	Business	Freshmen	Upperclassmen
E. *Things to see:*									
1. A famous movie star	or, 2. An historic site	(–)	(–)	(2)	(–)	(2)	(–)	(1)	(–)
1. A new type airplane	or, 2. A botanical garden	(1)	(1)	(1)	(2)	(1)	(1)	(1)	(1)
1. A chemical laboratory	or, 2. An army camp	(2)	(2)	(2)	(2)	(2)	(2)	(–)	(2)
1. An artist painting	or, 2. Animals in a zoo	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Men building a house	or, 2. Men grading tobacco	(1)	(1)	(1)	(2)	(2)	(1)	(1)	(1)
1. A giant bridge	or, 2. A large plantation	(1)	(1)	(–)	(2)	(2)	(1)	(1)	(1)
1. An old castle	or, 2. A giant motor	(2)	(1)	(–)	(1)	(1)	(1)	(2)	(–)
1. A new encyclopaedia	or, 2. A famous horse	(2)	(2)	(2)	(2)	(2)	(–)	(2)	(2)
1. A fashion show	or, 2. A science exhibit	(2)	(2)	(2)	(2)	(2)	(2)	(–)	(2)
1. A large crowd	or, 2. A great factory	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. Animals in the zoo	or, 2. A patriotic program	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. An army camp	or, 2. A giant motor	(2)	(–)	(–)	(–)	(–)	(–)	(2)	(2)
1. A science exhibit	or, 2. A large plantation	(1)	(1)	(1)	(–)	(2)	(1)	(1)	(1)
1. A dairy	or, 2. A library	(2)	(–)	(1)	(1)	(1)	(2)	(1)	(–)
1. An old castle	or, 2. A model farm	(1)	(1)	(–)	(2)	(2)	(1)	(1)	(1)
1. A great forest	or, 2. A business convention	(1)	(1)	(1)	(1)	(1)	(2)	(1)	(1)
1. A giant motor	or, 2. A novel filling station	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
1. An historic site	or, 2. A great forest	(–)	(1)	(2)	(2)	(2)	(1)	(2)	(–)
1. A business convention	or, 2. A science exhibit	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
1. A novel filling station	or, 2. An artist painting	(2)	(2)	(2)	(1)	(2)	(–)	(–)	(–)
1. An agricultural fair	or, 2. A pretty show window	(1)	(–)	(1)	(1)	(1)	(2)	(1)	(1)
1. A library	or, 2. A furniture factory	(2)	(2)	(–)	(–)	(–)	(1)	(2)	(–)
1. An auto show	or, 2. A herd of cattle	(1)	(1)	(1)	(2)	(–)	(1)	(1)	(1)
1. A flock of poultry	or, 2. A typewriter show	(2)	(2)	(1)	(1)	(1)	(–)	(–)	(2)
1. A model business office	or, 2. A model classroom	(1)	(1)	(1)	(1)	(2)	(1)	(1)	(1)

the other from the various pairs of choices presented in the test. Further comparisons were made between the interests of upper-classmen and a group of freshmen. By use of these varied comparisons a more complete analysis of the interests of a particular group of students is obtained and a fuller treatment of the nature of the students' interests presented.

1. *Comparisons of Interests of The Different Groups.* A comparison of choices made on the interest test by the different groups of students is presented in Table 2. An analysis of these choices reveals some interesting facts. In the occupational group of choices, the engineering students tended to select those that would be rated higher financially on an occupational scale than did the other groups. Their choices related to mechanical and manipulative work. Their interests in school subjects were to a very large degree in mathematics, sciences, and shop activities. These students would prefer to have about them objects which call for action in their use, such as a drawing set, a tool chest, and a radio set. Hence, most of their choices in "Things to Do" are of a manipulative nature, involving the use of hands and tools. These students desire to see scientific and mechanical work going on in industry. The nature of the work and requirements in the textile field are somewhat similar in nature to those of engineering. Many of the choices made by this group are similar to those made by the engineering students (see Table 3). However, their occupational choices revealed a greater interest in people and business as was shown by their choice of *merchant, lawyer,* and *high school teacher* rather than *electrician, telegraph operator,* or *bookkeeper.*

The forestry students' interests were found to be in occupations pertaining to forests and country life. The outstanding occupational choices, which differentiated them, were *forest ranger, cowboy, farmer,* and *farm management.* These students selected almost all of the school subjects which were related to that type of vocational choice, and desired to do things, as in the handling of domestic animals and out-door jobs found in the field of forestry. The interesting things that they wished to see related to forests, science exhibits, and large plantations. The

agricultural group of students indicated from their choices of interests that their interests were in activities connected with agricultural life, its problems and conditions. In their selection from the choices involving the varied school subjects, those subjects relating to agriculture were most commonly made. These students would rather have those things which are found in rural life; also, they would rather care for domestic animals, plant crops, and work with tools. It is rather hard to detect any very clear cut difference in the interests of these two groups. The forestry students are more inclined towards mechanical activities than are the agricultural students as indicated by their occupational choices. They chose *auto-mechanic, railroad engineer,* and an *electrician;* while the agricultural students chose an *express clerk,* a *college professor,* and a *merchant.* In "School Subjects," and "Things to Have," there were no differences discernible; but this greater mechanical interest on the part of the forestry students was again in evidence in the case of "Things to See."

The majority of the vocational education students were preparing to teach agriculture and their interests revealed a great deal of similarity with those of agriculture students; although they displayed a greater interest in business and people. In the selection of school subjects, they were found to be interested in vocational agriculture and other shop activities. Their interests showed a natural desire for domestic animals, a library, and shop tools.

In the business division, the interests chosen by the students revealed that they were interested in business activities. They made a fair selection of school subjects which would assist them in preparing for the business world. As future business prospects, they wanted to have, for example, *a new suit, a set of books, horses,* and *a typewriter.* They would rather see and do things which can be associated with the business field.

2. *Comparison of Response Patterns.* A comparison of the number of responses that were similar for the different groups is presented in Table 3. According to these comparisons students of vocational education and agriculture are most similar in interests, with textile and business students next and forestry and

TABLE 3

Comparison of the Number of Responses Similar for the Different Groups of Students

Divisions	Occupations	School Subjects	Things to Have	Things to Do	Things to See
Engineers and Textile	17	14	16	18	19
Engineers and Forestry	15	13	14	15	16
Engineers and Business	13	10	14	17	15
Engineers and Vocational Education	12	11	11	10	12
Engineers and Agriculture	10	13	11	11	11
Textile and Forestry	12	11	17	14	15
Textile and Business	21	12	17	24	16
Textile and Vocational Education	15	9	13	11	12
Textile and Agriculture	11	11	13	11	11
Forestry and Business	7	8	15	13	12
Forestry and Vocational Education	16	17	13	18	16
Forestry and Agriculture	22	18	13	21	14
Business and Vocational Education	15	8	12	10	9
Business and Agriculture	12	11	11	10	9
Vocational Education and Agriculture	23	17	14	20	17
Freshmen and Upperclassmen	10	13	18	20	16

agricultural students third. Business and agriculture students made the least number of similar responses with business and vocational education next and forestry and business third.

Thurstone (26) refers to eight reference factors that project into the various lines of endeavor. These factors are: (1) commercial, (2) legal, (3) athletic, (4) academic, (5) descriptive, (6) biological, (7) physical science, and (8) art. If we analyze the data of Table 2 in terms of these interest factors, it will be found that business would be classified as commercial, forestry as athletic, engineering as physical-science, agriculture as biological. Textile would be considered as a combination of the commercial and physical science, and vocational education would be classified under academic. Such a classification as presented by Thurstone of patterns is very useful, although these data indicate that there are innumerable patterns, and that many interests will reveal a rather distinct combination of two or more of these patterns. They do, however, tend to bear out the following statement from Thurstone (26) with reference to the physical science factor: "This interest factor differs distinctly from the biological group, and it seems necessary to conclude that the interests of college students in physical science and in biological science are rather distinct entities."

3. *Comparison of Upperclassmen and Freshmen.* A comparison of the upperclass students with the freshmen showed a degree of difference in interests in the list of "Occupations," and the list of "School Subjects," but a great similarity in their responses to "Things to Have," "Things to Do," and "Things to See." In thirty occupational choices, they showed similar responses in ten, dissimilar in eight, and non-differentiating in twelve. An analysis of these differences indicates a certain trend towards a commercial and legal interest factor. This is revealed in the upperclassmen's occupational choices of a *bank teller* or a *merchant* rather than an *electrician;* however, they chose an *electrician* to a *librarian.* They differed from the freshmen in that they chose a *store manager* over a *farm manager* and a *county farm agent* over a *match maker.* Yet, their responses, as shown by figures, displayed no degree of difference in interests in

the last three groups of the inventory. In "Things to Have," they both selected eighteen similar ones out of twenty, none dissimilar, and non-differentiating in only two. In "Things to Do," their similar choices were twenty, none dissimilar, and five non-differentiating out of a total of twenty-five. In twenty-five "Things to See," they chose sixteen similar interests, none dissimilar, and nine non-differentiating.

4. *Relationship to Environmental Background.* There is a difference of interests between the students with an agricultural home environment and those with an engineering and business home environment. The occupational interests of the students according to their respective home environments seemed to follow the field of work of their fathers. Those students whose fathers were in the agricultural field, showed a strong interest in the occupations pertaining to agriculture. The students whose fathers were either in the business or engineering field, showed a strong interest in the occupations pertaining to business or engineering, and very little toward agricultural interests. There were, out of thirty occupational choices, similar responses in twelve, twelve dissimilar, and six non-differentiating. In twenty interest choices of "School Subjects," these two groups of students responded alike in twelve, unlike in four, and with a negligible difference in four. They seemed to select those subjects which referred to their father's occupations. Yet, these two groups desired to have the same things as shown out of twenty choices, in which their responses were similar in fourteen, dissimilar in two, and non-differentiating in four. They showed only a slight difference of interest in "Things to Do," as the students gave similar responses in eleven, dissimilar in five, and non-differentiating in nine. In "Things to See," they differed greatly, having twelve like interests, six unlike interests, and seven non-differentiating interests.

Summary and Conclusions

The studies of interests that have been made during recent years indicate the growing application of more accurate procedures for measuring these more intangible phases of human personality. The applications of statistical techniques in the treat-

ment of the data gathered have shown certain fairly definite trends, characteristics, and relationships that should be of special value in various aspects of personnel work. The data for this study were gathered from the application of an interest inventory devised by the author and revised after trial experiences with college students. The major conclusions that may be drawn from this study of 320 students at North Carolina State College are:

1. The responses of the engineering and textile students indicated a marked similarity in the nature and direction of their interests. However, textile students were more interested in the business and commercial activities than were the engineering group.

2. Agriculture, forestry, and vocational education (largely agricultural) students were quite similar in their interests. It was significant that forestry students showed a greater interest in mechanical things, and that vocational education students were more interested in people and business than the agricultural group.

3. Business students displayed the business and commercial pattern of interests that one would expect of more mature workers in the business field.

4. In the comparison of the interests of freshmen with those of the senior college group a significant difference was noticed in "Occupations." The freshmen were interested in more specific types of occupations rather than those of a general and more business-like nature. No noticeable difference was present in the responses to the items listed as "School Subjects," "Things to Do," "Things to Have," and "Things to See."

5. There was a difference present in the choices of students from different home backgrounds. These interests indicate that the home has had a great influence in establishing interest patterns and trends.

6. The interest test here developed differentiates the interests of students in these technological fields. It seems quite likely that interest patterns appear among these students somewhat as Thurstone (26) has suggested and that such a test might be valuable in helping to find the nature of the pattern present in a particular student.

Bibliography

1. ADAMS, H. F. A scale for measuring psychological phenomena. *J. Appl. Psychol.*, 1933, **17**, 239–265.
2. ANDERSON, W. A. The occupational attitudes and choices of a group of college men. *Social Forces,* 1927, **6**, 467–473.
3. CARTER, H. D., PYLES, M. K., and BRETNALL, E. P. A comparative study of factors in vocational interest scores of high school boys. *J. Educ. Psychol.*, 1935, **26**, 81–98.
4. COWDERY, K. M. Measurement of professional attitudes. *J. Person. Res.*, 1926, **5**, 131–141.
5. COX, C. M. The early mental traits of three hundred geniuses. *Genetic Studies in Genius,* Vol. 2. Palo Alto, California: Stanford University Press, 1926.
6. DIMMICK, G. B. Differential interests of engineering graduates. *J. Appl. Psychol.*, 1936, **20**, 221–226.
7. FRYER, DOUGLAS. The objective and subjective measurement of interests—an acceptance-rejection theory. *J. Appl. Psychol.*, 1930, **14**, 549–556.
8. FRYER, DOUGLAS. *The Measurement of Interests.* New York: Henry Holt, 1931.
9. FRYER, DOUGLAS. Valuating measures of interests with particular reference to group interest scales. *Personnel Journal,* 1932, **11**, 103–110.
10. FURFEY, P. H. A revised scale for measuring developmental age in boys. *Child Develop.*, 1931, **2**, 102–114.
11. GOODFELLOW, L. D. A study of the interests and personality traits of prospective teachers. *Educ. Adm. and Supv.*, 1932, **18**, 649–658.
12. HUNT, THELMA. *Measurement in Psychology.* New York: Prentice-Hall, 1936. Pp. 187–204.
13. LEE, J. M., and SYMONDS, P. M. New type or objective tests: A summary of recent investigations. *J. Educ. Psychol.*, 1934, **25**, 161–178.
14. LEHMAN, H. C., and WITTY, P. A. One more study of permanence of interest. *J. Educ. Psychol.*, 1931, **22**, 481–492.
15. MINER, J. B. An aid to the analysis of vocational interests. *J. Educ. Res.*, 1922, **5**, 311–323.
16. NEMOITIN, B. O. Relation between interest and achievement. *J. Appl. Psychol.*, 1932, **16**, 59–73.
17. REMMERS, H. H. The measurement of interests difference between students of engineering and of agriculture. *J. Appl. Psychol.*, 1929, **13**, 105–119.
18. REMMERS, H. H. The validity, reliability and halo effect of human judgments in defined situations. *Psychol. Bull.*, 1933, **30**, 577.
19. STRONG, E. K., JR. *Vocational Interest Blank.* Palo Alto: Stanford University Press, 1928.
20. STRONG, E. K., JR. Diagnostic value of the vocational interest test. *Educ. Record,* 1929, **10**, 59–68.
21. STRONG, E. K., JR. Interest maturity. *Person. J.*, 1933, **12**, 79–90.
22. TERMAN, L. M., *et al.* Mental and physical traits of a thousand gifted children. *Genetic Studies of Genius,* Vol. 1. Palo Alto: Stanford University Press, 1925.
23. THORNDIKE, E. L. Early interests, their permanence and relation to ability. *Sch. and Soc.*, 1917, **5**, 178–179, and **6**, 359–360.
24. THORNDIKE, E. L. The interests of adults. II. The interrelations of adult interests. *J. Educ. Psychol.*, 1935, **26**, 497–507.

25. Thurstone, L. L. A multiple factor study of vocational interests. *Person. J.*, 1931, **10**, 198–205.
26. Thurstone, L. L. Factorial analysis of vocational interests. (A paper presented at the meeting of The American Psychological Association in Ann Arbor, September, 1935.)
27. Van Tuyl, K., and Eurich, A. C. Measuring the interests of college students with different major subjects. *J. Appl. Psychol.*, 1934, **18**, 27–44.
28. Wyman, J. B. The measurement of interest. *Voc. Guid. Mag.*, 1929, **8**, 54.
29. Young, Florence M. Causes for loss of interest in high school subjects as reported by 651 college students. *J. Educ. Res.*, 1932, **25**, 110–115.

THE RELATION OF INTELLIGENCE OF COLLEGE FRESHMEN TO PARENTAL OCCUPATION

B. F. Haught

THE UNIVERSITY OF NEW MEXICO

The question of the relationship of intelligence test score and occupational level has been the subject of a considerable amount of investigation. Most of the studies have reported positive correlation between these two factors, although there has been disagreement concerning the degree of relationship. There has been even more dispute over the question of the causal conditions which are operative in the determination both of 'intelligence' and of occupational level. The extent to which low cultural status is a cause or a consequence of low 'intelligence' is one of the major problems in the psychology of individual differences. The present study makes no pretense of attempting to solve this very difficult problem. It is possible, however, that a study of the relationship between these two variables in the case of college students would yield results which might have some bearing upon the general problem. Since most of the studies thus far have been carried out on grade-school children, it is conceivable that an analysis of the relationship in the case of college students would lead to different conclusions than those usually stated. For example, it might happen that the longer the exposure to a somewhat standardized American educational environment the more nearly equalized the individuals of different cultural backgrounds would become, and hence the lower the relationship between the test scores and occupational level. At any rate, the present report presents results secured from more than three thousand freshmen at the University of New Mexico, with a view to determining what this relationship might be.

Review of the Literature

Bridges and Coler (2), after examining 301 grade children, conclude that there is a striking correlation between the intelligence quotients and the occupational levels of the fathers. They even go so far as to say that if the mental ages were used as a basis for school entrance, children of the professional class should enter school two years earlier than those of the laborers.

Pressey and Ralston (9) in 1919 made a survey of 548 children in the third grade and found a positive correlation between occupational level of the fathers and the children's mental ability. There was, however, considerable overlapping of the groups.

Pressey (8) based a study on children below 14 years of age and found a positive correlation between paternal occupational levels and the child's intelligence, with children of professional and business men ranking much higher than children of the laboring class. He was inclined to conclude that these differences were due to variations in innate ability.

At Madison, Wisconsin, Dexter (4) made a study of 13 ward schools, including grades one to eight, and found an element of truth in the old saying "like father, like son." She discovered individuals in every vocation who rank very high in intelligence and others who rank very low. There was a tendency, however, toward a definite difference in intelligence according to occupational levels.

From data gathered in a rural survey for New York State, Haggerty and Nash (7) made a study of 8,121 school children in grades three to eight. The authors conclude that success in intelligence tests is in direct relation to the occupational level of the fathers, and that the latter is directly related to the probable success of the children in the American public school system.

Tonan Fukuda (5) reports a survey of the Washington School in Evanston, Illinois, dealing with 257 cases in the first eight grades and the kindergarten. He concludes that children whose parents are engaged in business or office work have decidedly higher intelligence quotients than do children whose parents are engaged in unskilled labor or peddling.

Terman (15) finds that 176 of his 560 'genii' come from the

professional group and only one from the laboring group. He concludes that these differences appear early and seem to be due to inheritance rather than to environment.

Sandiford (10) gave the modified Army Alpha to 5,052 high school, normal school and university students of British Columbia and concluded that children of parents in the professions had the highest median intelligence quotients. Then followed in order, the children of business and clerical workers, of skilled tradesmen, of farmers, of the semi-skilled, and of the unskilled.

Stoke (12) made a study from data gathered by himself and the Psycho-Educational Clinic of Harvard University of 508 children in the lower grades. He found a positive correlation between the occupational levels and the intelligence quotients.

From a study of 4,727 children in grades one to six in an Ohio school, Collins (3) concludes that the occupation of the father is a rough index to the intelligence of the child. He found that one-half of the children from professional parents were superior intellectually, and that not more than one-tenth of the children of unskilled laborers had a comparable degree of intelligence.

In a study of 380 children of pre-school age Goodenough (6) found that differences between the occupational groups are well established by the time the children reach the ages of two, three, and four years. These differences seem to be more dependent on innate factors than on the environment.

Stroud (13) administered the Pressey Classification Tests to 1,057 grammar school children and concluded that there is a low positive correlation between the intelligence of the child and the economic status of the parent when the latter is based on tax assessments.

Book in his volume entitled *Intelligence of High School Seniors* (1) finds more students from the professional groups than from the lower groups making high scores. From the first to the second level a gradual though not striking decline occurs in the scores.

Procedure

All freshmen of the University of New Mexico from 1921 to 1936, inclusive, comprise the subjects for this study. Starr has

analyzed the results from 1,730 of these cases, in an unpublished thesis (11). For the years 1921 to 1927 the Army Alpha test was used, and from 1928 to 1936 the American Council on Education Psychological Examination[1] was used. The scores for each test were converted into percentiles based on the University of New Mexico Freshmen, in order to have comparable scores for the two tests. On the census card which every student fills out is a place for the occupation of the father. From these cards the paternal occupations were obtained.

After the above information was procured, the next step was to find a suitable scheme for classifying the occupations and placing each occupation in the proper group. Taussig's classification (14), given below in abbreviated form, was adopted.

Group 1. *Unskilled.* Those comprising this group are the day laborers, who do their work through physical strength.

Group 2. *Semi-skilled.* Members of this group have specialized skill, bear some responsibility, and must have some alertness of mind. Examples are: railroad brakemen, policemen, barbers, mail carriers, truck drivers, chauffeurs, and factory hands.

Group 3. *Skilled.* Examples of this group are: masons, painters, carpenters, plumbers, shoemakers, machinists, mechanics, engineers, electricians, and blacksmiths.

Group 4. *Business and Clerical.* Examples are: foremen, office workers, salesmen, contractors, clerks, station agents, retail merchants, real estate workers, insurance agents, and manufacturers.

Group 5. *Professions.* A few examples of this group are: ministers, dentists, teachers, bankers, lawyers, brokers, druggists, accountants, editors, architects, and civil engineers.

Classifying the occupations was somewhat objective, although several difficulties arose. The original list included 4,342 students, but from this list 928 had to be excluded for the following reasons: (1) 328 gave the guardian's name instead of the father's; (2) 418 gave the mother's name, and in some cases her occupation, rather than the father's name and occupation; (3) 47 took the intelligence test, but failed to fill out the census cards and complete registration; (4) 103 failed to indicate on the census card the father's occupation; (5) 32 gave information which could not be regarded as indicating a vocation.

[1] Psychological Examination for High School Graduates and College Freshmen. Prepared by L. L. Thurstone. Published annually by the American Council on Education.

Results

The comparative data for the five occupational groups are presented in Tables 1 and 2, and in graphic form in Figure 1. Average percentile scores for the intelligence tests, and the corresponding standard deviations are shown in Table 1. The five groups exhibit a hierarchical order of intelligence test scores, corresponding to cultural level, with the highest test rating paralleling the highest cultural rating and so on down in descending order. The standard deviations of the five distributions are approximately equal—the lowest being 27.45 (professional group) and the highest being 29.23 (skilled labor group). The distribution tables, which are omitted in order to save space, show the range of scores to be about the same for each group. The

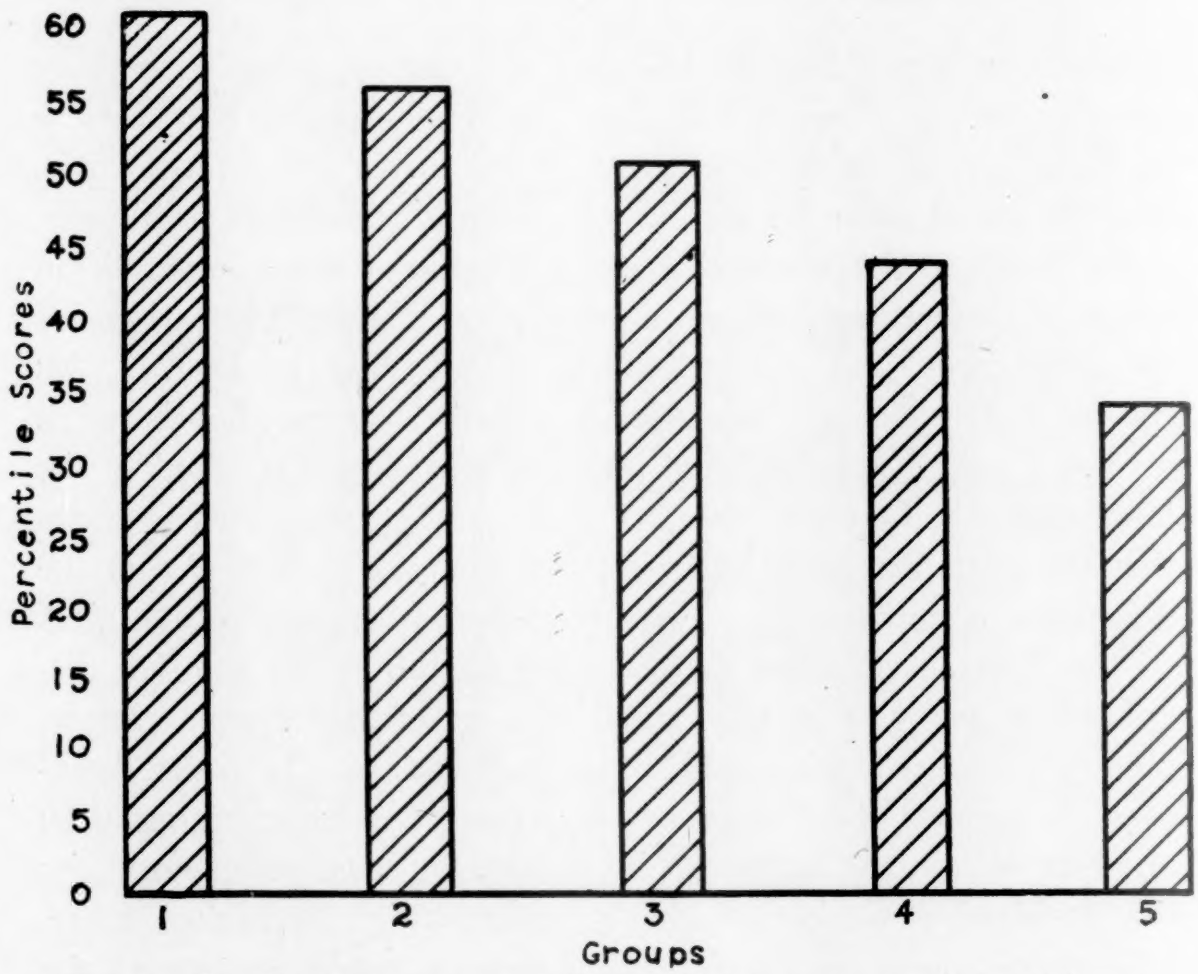


Figure 1. The average percentile score for each group as follows: (1) professional, (2) business and clerical, (3) skilled, (4) semi-skilled, (5) unskilled.

TABLE 1

MEANS, STANDARD DEVIATIONS AND NUMBER OF CASES FOR THE SEVERAL GROUPS

Group	Mean	Standard Deviation	Number of Cases
Professional	60.21	27.45	673
Business and Clerical	56.79	28.12	1,471
Skilled	50.38	29.23	439
Semi-skilled	44.89	28.93	748
Unskilled	34.72	27.65	83
Total			3,414

professional group and the business and clerical group have more cases at the upper end of the scale, the semi-skilled group and the unskilled group have more cases at the lower end of the scale, and the distribution for the skilled group is somewhat rectilinear.

In Table 2 the successive cultural groups are compared, in terms of differences between mean scores, standard deviations of the differences, and 'critical ratios' (Diff./S.D. of Diff.). All of the differences are statistically reliable, except that between groups 1 and 2 (professional and business classes). And in the latter case, the 'critical ratio' of 2.66 can be interpreted to mean that the statistical probability is 9,961 chances in 10,000 that the difference is a true one. With respect to the comparative sizes of the differences, since the standard deviations of the five distributions are very nearly equal, the gross differences can be directly compared, with only slight qualification. In order, however, to express the differences in terms of relative rather than absolute values, the latter can be divided, in each comparison, by the standard deviation of the group making the higher score. This procedure gives the following figures, corresponding to the groups compared in Table 2: .124, .228, .188, .35. The greatest

TABLE 2

RELIABILITY OF DIFFERENCES OF MEANS FOR SUCCESSIVE GROUPS

Group	Difference	S.D. of Difference	Diff./S.D. Diff.
M_1–M_2	3.42	1.287	2.66
M_2–M_3	6.41	1.576	4.07
M_3–M_4	5.49	1.751	3.08
M_4–M_5	10.17	3.214	3.16

amount of difference revealed is that between the semi-skilled and the unskilled groups; the smallest difference is between the means of groups 1 and 2 (professional and business). The comparisons for groups 2 and 3, and for groups 3 and 4 yield differences that are very nearly the same, in terms of the standard unit. Incidentally, the latter is to be interpreted in terms of the amount of surface (or percentage of cases) under the normal probability curve between any two means being compared. For example, the value .35, which represents the difference between the means of the semi-skilled and the unskilled groups, means that approximately 14 per cent of the cases lie between these two means. And this in turn means that 64 per cent of the semi-skilled group equal or surpass the mean of the unskilled group.

Conclusions

The results of these comparisons clearly exhibit reliable differences among the five occupational groups of college freshmen, in terms of intelligence test scores. The highest average test score is made by the group of the highest cultural (socio-economic) level, and a corresponding relationship holds for the other occupational classes and their respective average test scores.

But these figures do not reveal the cause, or causes, of these differences in mental test averages among the several occupational groups. Since the results for college students are found to be similar to those for grade school children, it can be concluded that the longer period of formal schooling of a somewhat standardized type does not equalize individuals of varying socio-economic backgrounds with respect to intelligence test performance. But this fact does not necessarily prove that such test differences as those observed here are due to "heredity," since the possibility remains that effective environmental conditions of "intelligence," other than "formal schooling," were operative. The advocates of the hypothesis of hereditary causation will still believe that individuals of the highest "native intelligence" are, on the average, drawn into the professions and that they transmit their capacities to their offspring—and so on for the other classes. The "environmentalists," on the other hand, will assert that

differences in cultural advantages and intellectual stimulation can account for the variations in average test ratings, since the latter depend upon success in responding to problems which involve specific cultural materials.

Bibliography

1. Book, W. F. *Intelligence of High School Seniors.* New York: Macmillan, 1929, 185–208.
2. Bridges, J. W., and Coler, L. E. The relation of intelligence to social status. *Psychol. Rev.*, 1917, **24**, 1–31.
3. Collins, J. E. Intelligence of school children and paternal occupation. *J. Educ. Psychol.*, 1928, **17**, 157–169.
4. Dexter, E. S. Relation between occupation of parent and intelligence of children. *Sch. & Soc.*, 1923, **17**, 612–614.
5. Fukuda, Tonan. Intelligence and environment of school children. *Amer. J. Psychol.*, 1925, **36**, 135–136.
6. Goodenough, F. L. Relation of intelligence of pre-school children to occupation of their fathers. *Amer. J. Psychol.*, 1928, **40**, 284–294.
7. Haggerty, M. E., and Nash, H. B. Mental capacity of children and paternal occupation. *J. Educ. Psychol.*, 1924, **19**, 559–572.
8. Pressey, L. W. The influence of inadequate schooling and poor environment upon results of intelligence tests. *J. Appl. Psychol.*, 1920, **4**, 91–96.
9. Pressey, S. L., and Ralston, R. Relation of general intelligence of school children to the occupation of their fathers. *J. Appl. Psychol.*, 1919, **3**, 366–373.
10. Sandiford, Peter. Paternal occupation and intelligence of offsprings. *Sch. & Soc.*, 1926, **23**, 117–119.
11. Starr, Mildred. Relation of intelligence of college freshmen to paternal occupation. An unpublished thesis filed in the library of the University of New Mexico, 1931.
12. Stoke, S. M. *Occupational Groups and Child Development.* Cambridge: Harvard University Press, 1927. Pp. 92.
13. Stroud, J. B. A study of relation of intelligence test scores of public school children to the economic status of their parents. *Ped. Sem.*, 1928, **35**, 105–111.
14. Taussig, F. W. *Principles of Economics.* New York: Macmillan, 1921, 134–137.
15. Terman, L. M. *Genetic Studies of Genius.* Stanford University Press, 1925, **I**, 61–72.

THE DESCRIPTIVE CATEGORIES OF PSYCHOLOGY [1]

LYLE H. LANIER
VASSAR COLLEGE

The Southern Society for Philosophy and Psychology has long seemed an enigma to persons not conversant with its peculiar mores. The 'alien' philosopher or psychologist finds it difficult to believe that this enlightened age would tolerate an institution designed to encourage the professional association of philosophers and psychologists. For have not psychologists spent almost a hundred years trying to escape the toils of philosophy? And have not philosophers, in turn, found scientific psychology so barren of intelligibility that they have fled to the orderly world of mathematical physics? Surely it must seem irrational and unrealistic in the extreme to perpetuate such a manifest anachronism as this Society! Indeed, certain of our own members have occasionally expressed embarrassment and chagrin, not merely that this illicit relationship should exist, but that it should be publicly proclaimed in the professional journals. The dissatisfied psychologist will insist that he is placed in a false light before his fellow scientists in other parts of the country; that whereas he is engaged in the tough-minded search for natural facts and laws, it might very naturally be suspected, on the basis of the principle of contiguity, that he trafficks in the fantastic laryngeal behavior constituting philosophical speculation; and, besides, that philosophers read such stultifyingly dull papers. Our philosophic malcontent, in turn, finds no inspiration whatever in the hunger drive of the white rat or in the learning of nonsense syllables; to say nothing of the fact that psychologists seem so crude and condescending, with their abundance of apparatus and paucity of ideas.

[1] Address of the president before the Southern Society for Philosophy and Psychology, Columbia, South Carolina, March 26, 1937. Certain portions of the original address have been expanded, in the interest of clearer exposition.

Fortunately, such disaffection has never been very pronounced in the Southern Society for Philosophy and Psychology, and I believe that in recent years it has virtually disappeared. Most of us are convinced that the relations between philosophy and psychology are intrinsic and fundamental, not merely historical and accidental. And, risking an accusation of provincial perversity, we should perhaps be willing to adopt as the official motto of the Society the following remark of the late Professor Hume, one of our charter members: " Philosophy without psychology is empty; psychology without philosophy is blind." This paraphrase of Kant's famous dictum does not imply that philosophy and psychology can be reduced to a common denominator. If any practice in the history of thought has produced more confused cerebration than psychologizing by philosophers, it has probably been philosophizing by psychologists. But there are legitimate and necessary relations between philosophy and psychology, and the effective cultivation of either subject requires an intelligent appreciation of these relations. It is especially important that such insight inform a discussion of the systematic concepts of psychology, and it will, therefore, be appropriate to outline briefly the general nature of these reciprocal interests before considering the special topic of this address.

The philosopher's interest in psychology differs from his interest in any other science, in at least two respects. In the first place, the essential terms of philosophic discourse must consist very largely of psychological concepts, since in the final analysis the psychological functions of man supply both the materials and the mechanisms of philosophical speculation. The definition of these concepts would seem, therefore, to be a very fundamental process in philosophy, and would require as precise knowledge of the genesis and development of psychological functions as scientific psychology could supply. Such concepts as 'sensation,' 'idea' and 'will' do not carry unequivocal connotations, even in contemporary psychology, and it is improbable that philosophers can supply valid 'existential' meanings for these terms simply by *a priori* proclamation. If philosophy would avoid a departure from whatever reality is represented in per-

ception and in scientific description, it would seem essential that the 'operational' meanings of terms current in scientific discourse serve at least as a foundation for whatever conceptual armamentarium philosophy might develop. Although it might reasonably be maintained that the ultimate goal of philosophy is to supplant the operational concepts of science with a body of existential or evaluative concepts, which would interpret adequately the flux of events constituting man's universe, the fact remains that the realization of this aim will not be promoted by the haphazard reification of loosely defined terms.

The second aspect of the philosopher's special interest in psychology relates to the nature of psychological 'knowledge.' Psychology inherited from philosophy an almost axiomatic belief in the *immediacy* of psychological knowledge, which was usually contrasted with the mediate character of knowledge in the physical sciences. All of the idealistic philosophies—whether rationalistic or mystical—would seem to be grounded upon such an assumption. But I shall try to show presently that the validity of this belief is now being seriously questioned by many psychologists, who consider that knowledge in scientific psychology is similar in its general logical character to knowledge in any other science. If this position be a sound one—and I am inclined to believe that it is—then those philosophers who prefer a 'psychological' approach to philosophical problems should attempt to define the status of the phenomena which give rise to their speculation. Perhaps it can be shown that a special form of knowledge, different in nature from the scientific variety, does in fact exist, and that this immediate intuitive insight is better suited to philosophical purposes. I am not interested in debating that issue; I would simply point out that philosophers of this persuasion should no longer beg the question.

With respect to the bearing of philosophical considerations upon psychology, it is becoming increasingly apparent that much of the confusion in systematic psychology has resulted from the failure to clarify the fundamental metaphysical assumptions which underlie psychological investigation. Often these basic postulates are not recognized; they are sometimes inappropriate;

they may even be contradictory, a condition illustrated by the use of dualism and monism concurrently as 'mind-body' postulates. Dr. Balz, in his presidential address last year (2), clearly showed the necessity for the elimination of the latter type of confusion from the intellectual architecture of scientific psychology. Occasionally, a psychologist may confuse metaphysical postulate with scientific hypothesis, as Hull apparently does when he suggests that only through rigid scientific procedure may one expect to arrive at a "valid conclusion concerning the ultimate nature of higher adaptive behavior"(7). Many scientists probably believe, as Hull apparently does, that the 'ultimate nature' of an event is to be defined solely in terms of the causal relations which it sustains with other events, and that the scientific method is the only way to determine what these relations are. But even upon the plane of this conception of 'ultimate nature' it must be recognized that one and the same event can participate in diverse types of relations, each of which might be investigated by a different science. Each science would render a highly abstract account of the causal relations revealed through the application of its methodology—the more accurate the account the greater the degree of abstraction—and the question naturally arises as to which of the scientific descriptions is to represent the 'ultimate nature' of the event. It might be maintained that the 'sum total' of the scientific descriptions would define the ultimate nature of the event, but then there is the difficult problem of deriving a basis for reducing these accounts to a totality, and of finding meaningful concepts for describing the synthetic product thus derived. All of which suggests that there is definitely a place for a philosophy of natural events which concerns itself with the *general* significance of all forms of scientific description, as well as with other possible types of 'knowing' or 'experiencing.'

A second major philosophical interest of the psychologist concerns the logical foundations of the processes by means of which he endeavors to establish a science of human nature. Specifically, the 'logic of scientific psychology' would involve an analysis of the activities performed by the psychologist in defining his con-

cepts, and in collecting, organizing and interpreting his data. Through such an analysis it is conceivable that valid criteria for resolving the conceptual chaos in systematic psychology can be established. Johnson (9) has suggested that modern psychologists have been at such pains to declare themselves emancipated from philosophy that they have at the same time freed themselves from the requirements of logical discourse. The consequences of such intellectual license have ranged all the way from illicit inferences from specific 'facts' to metaphysical systems concerning the 'ultimate nature' of psychological phenomena. Fortunately, in recent years the grounds of dispute in psychological theory have been gradually shifting from metaphysics to logical analysis, and we may reasonably expect that in time the formal outlines of a unified science of psychology will be defined.

The present paper will be concerned principally with the considerations which should govern the choice of the terms of discourse essential to a scientific description of psychological phenomena. I do not propose to enunciate a system of categories which would be expected to withstand logical attack and to provide for all future experimental developments. It is extremely improbable that a satisfactory repertoire of descriptive categories could be established either through *a priori* speculation, or through judicious selection from among the several systems of concepts which have been produced by the various 'schools.' An examination of the nature and basis of descriptive categories, with a view to developing a logically consistent attitude towards the problem, would seem to be a more profitable procedure at the present time. Such an examination will require the consideration of problems which lie within both of the spheres of philosophic interest ascribed to psychologists in the preceding paragraphs. I shall begin with the problem of the fundamental postulates of psychology.

The Postulates of Psychology

The term 'postulate' is used here to refer to the general presuppositions which underlie the scientific investigation of natural phenomena. Postulates, therefore, are general propositions con-

cerning the metaphysical nature of phenomena, and the validity of such assumptions is not subject to direct empirical test.[2] This being the case, the scientist should reduce his system of postulates to the absolute minimum required by the logical implications of his scientific activity, and by the self-evident character of the events investigated. At least two general postulates would seem to be universally accepted by scientists, and these may be stated as follows: first, that all events exhibit a lawfulness or rationality which can be discovered through systematic inquiry; second, that the world of natural events is objectively real. A third assumption, which most scientists would endorse, asserts that scientific explanations of events are to be found within the events themselves, not in extraneous agencies. Scientific explanation would then involve simply the determination of the fundamental laws governing the interrelations of events, and would be concerned neither with the question of existential reality, nor with the moot metaphysical problem of efficient causality.

I believe that most psychologists would accept these three postulates; but they disagree sharply concerning the assumptions which would define the specific nature of psychological events. On the one hand, there is the postulation of a fundamental duality in nature, in accordance with which psychological events are assumed to be qualitatively different from all other events. The antithetical postulate, natural monism, asserts that psychological events are continuous with all other events, and with the latter comprise the whole of nature. Unfortunately, there are no formal logical criteria whereby the scientist can evaluate the relative merits of two such postulates as dualism and monism, although the inability of scientists to utilize a postulate as a basis for successful scientific activity would constitute evidence strongly presumptive of the postulate's worthlessness. By this test of efficacy in promoting the development of an organized body of knowledge, it is my firm conviction that monism will prove to be far more serviceable than dualism as a foundation for psycho-

[2] This definition of postulate should be distinguished from Hull's usage in a recent paper (7). I would use the term 'hypothesis' to refer to the type of proposition to which Hull applies the term 'postulate.'

logical investigation. In support of this belief, one can adduce, negatively, the rather complete failure of the dualistic psychology of consciousness to develop into a systematic science. An important positive argument for monism can be found in the fact that representatives of diverse systematic points of view, although often disagreeing with respect to specific descriptive concepts, nevertheless insist upon the desirability of a 'monistic' psychology (**3**, **6**, **12**, **14**).

Only a casual acquaintance with the status of contemporary psychology will reveal, however, that the generality of psychologists are far from consistent in their attitudes toward the 'mind-body' problem. As suggested above, the practice of postulating both dualism and monism *implicitly* is not uncommon, the choice of the postulate varying with the type of problem under consideration. I shall not digress to view this practice with alarm, nor to burden you with illustrations which flourish in most of the psychology text-books. Instead, I should like to raise the question of why psychologists should be so peculiarly subject to this type of intellectual confusion. I refuse to believe that psychologists are less well endowed, as a scientific body, than scientists in other fields. I prefer to think that the difficulty lies mainly with the manner in which most protagonists of monism have elected to utilize this assumption in defining the subject matter and in developing the systematic concepts of psychology. The so-called objective psychologists have had, until very recently, a virtual monopoly of monism, and their behavior in respect of this principle has been such as to antagonize, confuse, or frighten the generality of psychologists. Many of the latter group, while admitting the futility of the 'psychology of consciousness,' nevertheless have felt that the older psychology was based upon concrete observable phenomena, and that to disregard these phenomena was neither good psychology nor good sense. Even such a philosophically sophisticated statement as Lashley's 'behavioristic interpretation of consciousness'(**14**) apparently could not allay the suspicions regarding the adequacy of a thorough-going monistic psychology.

The specific practice which has militated most strongly against

the acceptance of monism has probably been the objectivist's obsession for translating psychological events rather literally into anatomical and physiological events. The objectivist, to be sure, has insisted upon a difference between the behavioral events which he investigates and the events of the physiologist; the behavioral events have usually been described as complex activities of the organism as a whole, whereas the physiological events have been considered to be activities of the parts of the body. But this formal distinction has not prevented the creation in psychology of a physiological mythology which often is as fantastic as the older mentalistic mythology. The behaviorist presumably has excused the new mythology on the grounds that the close association of his somewhat indeterminate subject with the solid substance of anatomy and physiology would render his discipline more scientific. Now monism obviously implies that behavioral or psychological events are organic in nature, but it does not follow that these events can be adequately described in terms of the scientific concepts of anatomy and physiology. The objectivist, as well as many of his critics, has often been guilty of an elementary fallacy of equivocation in the use of such terms as physical and physiological, and this practice has undoubtedly been the source of much antagonism towards the general monistic point of view. The term physiological, for example, may be used to refer to the general fact that an event is an exhibition of organic activity, quite irrespective of what is known about the nature of the activity; but on the other hand, 'physiological' may refer to knowledge as represented by the concepts of the science of physiology. It is clear that the postulation of monism commits the psychologist to the general position implied in the first meaning of the term, and therefore that he need not try to prove, by tortuous references to hypothetical activities of nerve, muscle or gland, what he has already assumed as a basic presupposition of his work. But the monistic postulate by no means implies that psychological events are to be explained (*i.e.*, described scientifically) in terms of the concepts and methods which define the science of physiology.

In view of the elementalistic bias of most monistic psychology,

it seems likely that the adoption of a fifth postulate would resolve many of the doubts concerning the adequacy of monism. This postulate would assert the existence of different levels of organization within nature, and would deny that the phenomenal properties characteristic of any level can be explained, without residue, in terms of the concepts appropriate to a lower level. Accordingly, the belief in the fundamental continuity of natural events would not imply that human behavior can be described relevantly in terms of the scientific constructs of physiology, of chemistry, or of physics. Not that every aspect of human behavior might not be studied by physicists, or by chemists, or by physiologists; but their concepts would not adequately represent the essential characteristics of behavioral events. This principle would not deny that knowledge concerning events at lower levels, especially at the level of physiology, could reveal important conditions of psychological activities, nor that physiological data might often provide suggestive hypotheses for the guidance of experimental investigation in psychology. And likewise, the understanding of the more complex psychological phenomena would undoubtedly be facilitated by the study of the manner in which they condition and are in turn affected by the level of events comprised within the social sciences. There is no *a priori* reason why explanation in terms of the concepts of a more complex level of natural phenomena should not be just as revealing as explanation in terms of the analytical constructs of a simpler order of events. But neither type of explanation will be directly relevant to the range of events which are denoted by the term 'psychological.' Psychology must develop its own distinctive concepts if it proposes to establish an adequate body of knowledge concerning psychological events.

The General Nature of Psychological Events

The term 'psychological event' has been frequently used in the foregoing pages without explicit definition. As a matter of charity, we have assumed that psychologists do investigate a distinctive order of phenomena and that these phenomena are neither physiological events nor sociological events. The lines of

demarcation between psychology and these two related sciences may be indistinguishable at the respective margins, but in self-defense psychologists must proclaim the existence of an intervening order of natural events, the understanding of which requires the special methods and conceptions of scientific psychology. Beyond this gesture of self-preservation, however, unanimity of opinion among psychologists disappears. The peculiar development of 'schools' of psychology reflects the extent of the failure to arrive at uniform conceptions concerning either the nature of psychological events, or desirable methods of investigation. The varied claims of behaviorism, renovated structuralism, configurationism, purposivism, and personalistics—not to mention such choice by-products as psychoanalysis and parapsychology—must surely suggest to the layman that psychology is in fact a scientific tower of Babel.

The general definition of psychology as the science of human behavior, or as the science of the behavior of persons, should be sufficiently catholic in connotation to include all of the behavioral properties characteristic of the natural system which we designate as 'man.' Psychologists as widely divergent in their systematic points of view as McDougall, Koffka and Watson have all used the term human behavior to denote the general subject matter of psychology. The fact that different psychologists may approach this subject matter with somewhat specialized theories concerning its nature, and with interest in selected portions of the totality of behavioral events, does not impair the general usefulness of the concept. For the term human behavior refers not merely to overt motor response, but also to such characteristic human activities as perceiving, imaging, wishing and thinking. We shall consider presently whether or not these are valid scientific categories; I use the terms here in their commonsense connotations, in order to indicate that "behavior" need not be defined as mere muscle-twitching.

I should like to emphasize the implications of the statement in the preceding paragraph that different psychologists may approach behavioral phenomena with special interests in narrow portions of the total range of events. If one can presume to

generalize in respect of matters so complexly conditioned, it might be said that most of the 'schools' of psychology have originated through too narrow preoccupation with events of a given type, and through the false assumption that the descriptive concepts derived from the study of such events are applicable directly to events representing all other levels of behavior.[3] Sanborn has written very pertinently (17) that "the method must accommodate itself closely to the subject-matter as found empirically and . . . interpretative concepts must establish their value experimentally." In psychology, "the subject-matter as found empirically" would seem to range all the way from simple sensory discrimination to the manifestations of individuality referred to by the term 'personality.' The methods and concepts which are adequate to the study of the first problem will be found to be irrelevant as applied to the second one. In general, an eclectic attitude towards the sensible methodological approaches would seem to be appropriate, but the eclecticism which I recommend does not imply that such methologies would be logically equivalent as applied to each type of problem, nor that the rendition of the facts by one 'school' can be translated without alteration of meaning into the vernacular of any other 'school.' Instead, it follows from the fifth postulate laid down above that there should perhaps be developed an hierarchy of methodological concepts corresponding to such phenomenal levels of behavior organization as might be discovered.

Any discussion of the nature of psychological events must inevitably face the issue of whether or not behavioral data are both objective and subjective. Or, to express the question in more conventional form, does not psychology study both behavior and experience? If psychologists could arrive at common agreement concerning this question, on the basis of the precepts of the universal logic of science, much of the intellectual confusion in systematic psychology would disappear. Naturally, the possi-

[3] Essentially this same idea has been expressed by Rosenzweig (16), in an article which I read after this paper was written. He uses the phrase 'fallacy of arrogation' to describe the practice of imposing upon the totality of behavioral events the descriptive concepts developed in relation to a special class of activities.

bility of arriving at agreement depends first of all upon unanimity in the choice of a 'mind-body' postulate, and since most psychologists subscribe verbally to a monistic position, we shall proceed with the discussion upon the basis of this general presupposition. It should be noted, however, that although the postulation of monism eliminates from consideration the concept of mind or consciousness as an extra-organic process, at the same time it is not logically inconsistent with the assumption of levels of organization in nature that a novel property should emerge which would have the characteristics conventionally ascribed to consciousness, or 'direct experience.' It is not the business of the psychologist to sit in judgment upon nature, to require that she exhibit only those properties which can be conveniently investigated. In order to be scientific it is not necessary to distort phenomena by moulding them into an arbitrary and over-simplified conceptual pattern. The psychologist must accept with a considerable degree of naïvete the world of psychological events as they are presented to him. He has only one limitation, or perhaps it is his principal advantage: his account of these events must be a scientific account, not a poetic or a philosophical account. Let us examine the problem of consciousness in the light of this injunction.

The conception of consciousness as a unique order of events which comprises the distinctive subject matter of psychology was scarcely questioned at all prior to the beginning of the present century. Irrespective of differences in their approaches to the study of consciousness, or in their conclusions regarding the interrelations of conscious phenomena, the earlier psychologists were almost unanimous in the belief that the business of psychology was to investigate 'immediate experience.' The psychologist might concentrate upon conscious content, or upon conscious acts, or upon the stream of consciousness, depending upon his systematic position with reference to the most appropriate way of describing consciousness, but to have doubted either the reality or the relevance of consciousness for psychology would probably have been considered indicative of some form of insanity. The notion of psychology without consciousness

would have seemed as ridiculous as that of anthropology without man. But in 1937 it is not at all unusual to read definitions of psychology which make no mention of consciousness or immediate experience, and which nevertheless are propounded by serious and reasonable men who purport to include within their definitions all of the distinctive psychological phenomena which are open to scientific study. And even when the term 'consciousness' is used, it is offered, as with Boring[4] (**3**), within a logical framework which serves to differentiate the concept completely from the usage current in the older psychologies. The history of this radical transformation is too involved for detailed treatment in this paper, but it will be instructive for the study of psychological categories to consider just what the 'loss of consciousness' in modern psychology implies. Especially so, inasmuch as many psychologists still feel that in the process of losing consciousness the psychological baby has been let out with the bath.

In the psychologies of consciousness there were two essential aspects which should be carefully distinguished: (1) consciousness, or immediate experience (conscious states, conscious activities, direct experience), and (2) introspection (the means by which knowledge about consciousness was secured). Let us consider the first concept, by asking the following question: how is consciousness to be defined? If one receives the naïve answer that the phenomena denoted by this term cannot be defined, but can only be experienced directly, then one must perforce conclude that such material is not subject to scientific enquiry. Lashley has observed (**14**), correctly it seems to me, that an event which can be known only in terms of itself, which cannot be exhibited or represented in terms which permit it to be related to other events, is by such definition automatically removed from the province of science. Unique qualitative essences which can only be experienced privately must remain with the individual since he will find no way to render them socially intelligible. Alex-

[4] In a recent article (A Psychological Function is the Relation of Successive Differentiations of Events in the Organism, *Psychol. Rev.*, 1937, **44**, 445–461) Boring presents an exceptionally clear analysis of the difficulties involved in the definition and use of the concept "consciousness," and recommends that it be dropped from psychology.

ander (1) has tried to preserve the intimate immediacy of experience, while denying its validity as knowledge, by distinguishing two aspects of experience, which he calls contemplation and enjoyment. Contemplation refers to ordinary observation or perception, which yields our primary knowledge of ourselves and of the external world. But contemplation is also enjoyed, although enjoyment can never become a datum of knowledge. I recommend this solution to all psychologists who would eat the cake of experience and yet have it.

It should be obvious that the introspective psychologists could not really have subscribed to the conception of consciousness as an ineffable and unknowable process, else they would have been compelled to relinquish their subject to the yogi. They believed instead that by introspective analysis they could arrive at quite systematic knowledge concerning consciousness. They performed innumerable experiments, made elaborate reports concerning the contents of consciousness, and developed systematic theories about the relations of these contents to the whole of consciousness and to mind. In spite of the devastating criticism which has been levelled against the many vulnerable aspects of structural psychology, it must, I think, be admitted that its protagonists uncovered a great body of factual information. But it seems equally clear that their systematic tenets caused them to misrepresent the character of this information. They believed that their reports constituted scientific descriptions of observed aspects of consciousness, but an examination of their work reveals that they usually were confusing facts of assumption with facts of observation. Johnson (10, 11) has presented a detailed analysis of the nature of this confusion in the case of the measurement of sensory thresholds, a field in which most introspectional psychologists believed that precise descriptions of immediate experience had been accomplished. Johnson showed clearly that 'Fechnerian sensation' was a construct which derived its meaning from a specified set of mathematical and experimental operations, and that 'introspectional sensation' had never been shown to be identical with or even related to the Fechnerian concept. In general, I believe that it is now recognized by most

introspectional psychologists that instead of revealing or measuring consciousness they were principally engaged in the study of sensory discrimination. Boring (3) is perhaps the outstanding exponent of this general point of view, and he has recently gone so far as to define consciousness in terms of the specific discriminative responses of the organism. It is his view that discrimination (consciousness) can be studied analytically in terms of certain dimensions, each of which is presumably susceptible of definition in operational terms. Whether or not Boring's four dimensions provide adequate analytical constructs for describing all of the observable phenomena conventionally subsumed under the concept consciousness, might be subject to debate, especially on the part of Gestalt psychologists. But there can be no doubt that, in its formal outlines, Boring's approach defines a basis for the scientific investigation of all of the introspectionist's phenomena which are susceptible to systematic observation. Under the exegesis of operational logic, the uncertainty and indefiniteness which characterized the futile efforts to describe the mentalistic entity 'consciousness' should disappear.

With the realization of this desideratum, the principal difficulty with the notion of introspection would be removed. In the schema of structural psychology introspection was conceived as a sort of 'consciousness of consciousness'—a process whereby immediate experience was observed—and this notion was so obscure that even the introspectionists themselves could do little more towards clarifying it than to exhibit their protocols as evidence that the process actually worked. But we need now no longer be disturbed about the conceptual status of introspection defined as an inner sense through which experience is revealed. The essential fact concerning introspection would seem to be that the individual *knows* about his past behavior. I quite subscribe to Boring's view (3) that introspection is the act of *remembering* what one has seen, done, felt, thought, etc. The process of knowing, or remembering, is indeed unusual, but the mystery is certainly not lessened by the postulation of nebulous psychic processes called consciousness and introspection. The principle of parsimony would seem to dictate the avoidence of an explana-

tion of the unknown in terms of the still more unknown. The concept of remembering implies that some definite and orderly means of symbolizing past activity exists and that under appropriate conditions the individual can represent past behavioral events through some form of report (verbal, graphic, etc.). All types of behavioral processes, irrespective of locus of crucial stimulating conditions, are assumed to be logically correlative with respect to the possibility of symbolic representation, and hence of subsequent recollection. The distinction between objective and subjective observation would seem then to be mainly dependent upon differential physical and physiological conditions rather than upon a fundamental qualitative diversity in respect of the nature of the knowing process. In consequence of the variations inherent in the anatomical and physiological processes involved, there would be differences in the extent and in the accuracy of remembering such varied behavioral events as, for example, intra-organic discrimination, extra-organic discrimination, emotional reactions, thinking. But the distinction between mediate and immediate knowledge could with great profit be banned from psychology, in so far as this distinction implies that psychology has a special avenue of approach to knowledge of its subject-matter which is denied to other sciences. A better statement of the situation would be that a great deal of psychological fact involves reports (introspection, remembering, observation, or what not) concerning behavioral events which depend almost entirely upon intra-organic conditions, whereas all other sciences deal with reports involving behavior with reference to extra-organic conditions (cf. Köhler, **13**, pp. 19–34). In *all* instances the *knowledge* derived from scientific observation is essentially mediate or constructural in nature, not immediate or intuitive.

The Problem of Descriptive Categories

The definition of psychology as the science of human behavior, or as the science of the behavior of persons, serves to delimit a general field of natural events, but since these events exhibit quite diverse phenomenal properties the psychologist must seek to develop corresponding conceptual distinctions. Commonsense

observation had produced a great variety of descriptive concepts long before psychology emerged as a special science. In fact, this type of observation not merely distinguished many different aspects of behavior but often accentuated the degree of such differences through the failure to detect underlying similarities of pattern or of causal relation. Science attempts to simplify the understanding of the flux of events which present themselves to the casual observer, both by reducing to a minimum the number of descriptive concepts and by expressing causal relations in terms of the fewest possible laws. But this precept of maximum simplification, or parsimoniousness, may lead to a gross distortion of reality, through the failure to recognize the relevance, for scientific description, of concepts which represent different levels of phenomenal organization. This form of scientific malpractice is best illustrated in psychology by the attempts to use a few atomistic constructs, such as sensation or conditioned reflex, to explain the entire range of psychological phenomena. The psychologist, it would seem, must move cautiously between the Scylla of an over-simplified conceptual schema and the Charybdis of a vocabulary which represents no advance over the uncritical language of the man in the street.

Recent years have witnessed a growing dissatisfaction on the part of psychologists with the systematic analysis of psychological events represented by the traditional categories. Hunter, for example, has remarked (8) that psychology is the only science which has not been able to continue with the subject matter bequeathed to it by philosophy, and in accordance with this belief he suggests that the entire system of classification developed by philosophy and the older psychology be discarded. Included under Hunter's ban would be the following concepts: "those major aspects of mind, the Intellect, the Feelings and the Will, with minor distinctions of sensation, image, memory, judgment, thought, simple feelings, emotions and conations." Hunter admits that objective psychologists have provided no satisfactory substitute for the older systematic psychology, and asserts that the development of appropriate concepts is one of the main problems of contemporary psychology. He proposes

that psychologists devote themselves entirely to the study of the extrinsic behavior of the organism in its relation to the social environment, particularly to such general classes of behavior as the following: (1) behavior aroused by stimulations in which other persons are integral parts, such as jealousy, anger and rivalry as an incentive for work; (2) behavior whose results are socially important, such as type-writing, home-building and mathematical behavior, (3) behavior which serves as a tool for social conduct, such as language activities, and (4) behavior maladjustments. Inasmuch as Hunter does not attempt to translate the phenomena comprised within these general areas of investigation into specific behavioral concepts, it is difficult to evaluate his approach in terms of its systematic possibilities. His primary concern would seem to be that mentalism be banished from the systematic architecture of scientific psychology, and with this impulse I am in hearty accord. Likewise he would seem to be on sound ground in insisting that psychologists should derive their concepts from phenomenal behavior rather than from hypothetical physiological processes, although his term 'extrinsic behavior' is perhaps an inadequate rubric for designating the totality of psychological events which are susceptible of scientific study.

It is to Tolman (20) that one must turn for the most comprehensive effort to describe psychological events in terms of systematic objective constructs. Like Hunter, Tolman believes that the descriptive concepts of psychology should be derived directly from behavioral (molar) processes as such, rather than from assumed physiological (molecular) processes. And, again like Hunter, Tolman proposes that psychological concepts be divested of all mentalistic connotations. But instead of disregarding the phenomena to which the mentalistic terms have been unfortunately attached, as Hunter is inclined to do, Tolman elects to rename the concepts, or else to redefine them in unambiguous behavioral terms. The contrast between Tolman's program and that of the earlier ('stricter'?) behaviorists is well illustrated by the following quotation:

"But, after all, we really cannot escape the old question of *sensation* and *image*, of *feeling* and *emotion*. The good old psychologists in their laboratories, who introspected and filled innumerable pages of their Protokolls with accounts of these processes, were doing something and doing it ably. What, now, in *our* terms, was this that they were doing?" (**20**, p. 234).

He then proceeds in operational fashion to define these 'subjective' phenomena, and although many psychologists might disagree with him concerning the meanings assigned to specific concepts (e.g. those relating to feeling and emotion), the logical basis of his procedure would seem to be sound.

Inasmuch as we are interested in the formal aspects of the behavior of concept-makers, rather than in the validity of particular concepts, an examination of the general features of Tolman's system should be valuable for our enquiry. In the first place, it should be noted that he by no means restricts his efforts to mere renaming or redefining of traditional concepts, as the preceding paragraph might have seemed to imply. On the contrary, he attempts a constructive reformulation of the conceptual framework of psychology, in terms of concepts derived directly from behavior. Tolman's general mode of procedure may be roughly outlined as follows: (1) a description of the universal characteristics of behavior, as a general type of natural event; (2) the establishment of the initiating causes of behavior; (3) the identification of the major classes of behavior determinants, other than the 'initiating causes'; (4) the elucidation of the general principles developed in (1), (2), and (3), with respect to the behavior of the rat in the maze (involving the use of such descriptive categories as seemed to be required by the experimental facts); (5) the elaboration of the conceptual repertoire to the extent necessary to encompass the progressively more complex behavior of the 'higher' animals, including man. Whether or not Tolman's system, *in toto*, will meet with widespread 'social' acceptance,[5] his method of deriving concepts exhibits at least three very important features. First, he accords priority—both logical and empirical—to the descriptive concepts which rep-

[5] The 'awkward' terminology, especially the many hyphenated words, will perhaps militate against the assimilation into ordinary psychological discourse of many of Tolman's concept-names.

resent the most general and inclusive relations involved in behavioral dynamics. Second, after thus envisaging a broad organismic foundation for his system he proceeds, in effect, from "general to particular" by a sort of progressive analysis of behavioral manifestations. But in this process the special aspects of behavior, which are represented by corresponding inferential constructs, are conceived in the light of their basic relation to the total behavioral situation. Thus Tolman would seem to have obviated the possible criticism that his concepts are products of what Lewin (**15**) has called abstractive classification. A third feature of the system is the intimate relationship which prevails between his concepts and experimental facts; his categories emerge from observation and tend to maintain an operational relation to scientifically observed phenomena.

Tolman's system can be said to represent behaviorism grown mature and philosophically sophisticated. It is interesting to observe the unusual degree of logical conformity between this system and that of Lewin (**15**), which is a corresponding outgrowth of Gestalt psychology. Lewin, like Tolman, is concerned primarily with 'molar' behavior, and denies the relevance, for psychological description, of physiological processes and theories (**15**, pp. 79–80). Behavior is conceived to be a natural phenomenon in its own right, and should be investigated in terms of concepts capable of representing its dynamic (causal) relations. In order to comprehend the totality of these relations within a unified conceptual system, Lewin proposes that the basic concepts of psychology should represent the person as a whole in dynamic interaction with a total environment. To the total psychological situation thus conceived Lewin gives the name 'psychological life space,' and this concept provides a basis for the derivation of all psychological events which might possibly occur under a given set of conditions. The progression from general to particular—which is implicit in Tolman's work—is an explicit creed with Lewin. The latter is interested in establishing a formal method whereby the facts in all of the fields of psychology can be brought within the purview of a single system of concepts, which would itself constitute a sort of Gestalt.

Tolman works empirically with experimental behavioral facts, beginning with the generalized behavior of the rat and progressively elaborating his conceptual armamentarium to include the highly specialized behavior of the human adult. Lewin is not so much concerned with the systematic organization of existing psychological fact, per se, as he is with the development of a formal conceptual mechanism whereby all possible behavioral data might be represented. In order to avoid the limitations which attach to existing concepts—they are said to be products of 'abstractive classification'—Lewin selects the branch of non-metrical mathematics, topology, as a conceptual medium. Topological psychology may be defined as a system of spatial schemata which represents the relations that can exist between a person and his environment—particularly whole-part relationships and the relationships between the parts. But since the topological schemata represent all possible events within a given life space, a complementary 'vector' psychology will be necessary to determine which events actually occur in a given case. Thus far the vector concepts have not been developed systematically in relation to the experimental facts of psychology, and until this is done it will be impossible to determine the efficacy of Lewin's system.

Meantime, psychologists must continue their research, and this research must be guided by concepts which give promise of pragmatic acquaintance with behavioral reality. Lewin claims that the existing concepts of psychology are valueless, since they imply static behavioral entities which usually are explained merely by reference to the properties with which they are endowed by definition. He seems to imply that the investigation of the specialized aspects of behavior must await the development of knowledge concerning the more complex phenomena of "will, needs and personality." This conclusion would seem to follow from his belief that the specialized concepts should be derived from the more general ones—or in relation to the latter—by a procedure of 'gradual approximation.' This method would insure 'right simplification,' he thinks, whereas the illicit process of 'abstractive classification' produces isolated categories and a

merely statistical type of investigation. Most psychologists would readily acknowledge that many of the traditional categories are as bad as Lewin claims and they would naturally welcome a method which would produce a fool-proof system of 'right' concepts. But I cannot believe that a final hierarchical configuration of concepts, which any science would acclaim, must inevitably be built up according to this formula of "progression from general to particular." This has scarcely been the course of events in the more mature natural sciences—not even in physics, a field in which Lewin finds the most complete departure from the method of abstractive classification toward that of constructive generalization. The science of physics did not arrive at its present state by beginning the study of the behavior of inanimate objects in terms of all-embracing concepts which proposed to envisage all forms of energy relationships simultaneously. Instead, the physicist has distinguished several forms of such behavior, each of which requires its peculiar concepts and special experimental methods. If the investigator is interested in the conditions of motion in macroscopic objects, he develops the methodological concepts which prevail in the general field of mechanics. Other events will be subsumed under such categories as heat, light, electricity, and so on. It may well be the case that all of these types of energy manifestations depend upon an underlying unitary order of events, but this fact would not invalidate the several descriptive categories, with their special concepts, as proper methodological tools for their respective forms of organization. I am convinced that the same logic must prevail in psychology. And while the subject-matter of psychology—the behavior of a somewhat unitary person—is perhaps relatively more homogeneous than that of physics, nevertheless, the phenomenally diverse aspects of behavior must be differentiated in terms of operationally relevant concepts, in order that appropriate hypotheses and methods of investigation can be developed.

It is obvious, of course, that the psychologist is much less certain about his categories than is the physicist, and Lewin has exhibited certain of the important causes of this state of affairs. The traditional categories which psychology inherited from com-

mon-sense observation and from philosophy have too often been conceived as discrete faculties endowed with existential properties. Investigation on the basis of such a conceptual framework would naturally tend towards mere classification, or identification of the faculties believed to be operative in a given behavioral situation. Another way of expressing a consequence of this unfortunate type of procedure is to say the psychologists concerned have lost sight of the fundamental objective of science, which can be defined as the determination of cause-and-effect relationship among natural events. The problem is to explain the behavior and not to proclaim the behavior to be self-explanatory. Lewin would certainly seem to be on sound ground, therefore, when he rejects all behavioral categories which are conceived out of relation to the situations in which presumably they are believed to function. But this rejection really is equivalent to saying that such categories have not been operationally defined (**11**, **19**) but have instead been conceived as existential entities. The requirement that the descriptive concepts of psychology have operational meaning and that the use of concepts be motivated by the basic scientific purpose of deriving causal laws would cure most of the ills of systematic psychology. And, contrary to Hunter's belief, such treatment would not mean that the traditional categories would have to be discarded *in toto*. Certain of these terms, such as instinct and sensation, have unfortunate systematic connotations, but the latter concept can be defined operationally to denote the discriminative behavior which occurs under certain conditions and is certainly as good as Hunter's term receptor process. In the case of instinct we have a concept which has not been shown to be definable in terms of specific operations, and it has no place in scientific psychology. Such concepts as perceiving, learning, remembering, imaging, thinking, feeling and emotion can probably be defined in terms of distinctive sets of operations and there would seem to be no particular advantage to discarding them because of past systematic abuses.

English (**5**) has objected to most of these descriptive categories on the ground that they have unity only for the observer, not in terms of the inherent character of the acts themselves. The

category 'learning,' for example, includes a great variety of phenomena which seems to be only very superficially related among themselves, and English would abandon the concept in favor of a search for genuine functional unities of behavior. He would begin with practical problems of everyday life and through careful experimentation and statistical treatment of results would expect eventually to discover the desired functional unities. The latter would be found, he believes, somewhere between the common-sense units and the constituent physiological processes. English's criterion of a functional unity is not very clearly defined, but to some extent he would seem to be seeking a new assortment of existential entities of the nature of fixed types of behavioral patterns. Irrespective of the accuracy of this interpretation of English's position, it is becoming increasingly clear that psychologists must avoid any criterion of the validity of concepts which would imply the existence of fixed, mutually exclusive entities—whether behavioral or physiological. Since all descriptive categories represent abstractions from the totality of the behavioral events which are characteristic of a somewhat unitary person, there will of necessity be great overlapping with respect to the specific events comprised within these categories. The differentiation of concepts will depend upon the possibility of denoting the diverse forms of relations which a person can sustain with his environment in terms of correspondingly diverse patterns of operations. Such general categories as those mentioned above imply that it is possible to denote a broad class of person-environment relations by means of operations which exhibit close similarity, or identity, in respect of the *general form of the experimental or mathematical procedures* which constitute these operations. Within each of these general descriptive categories, such special concepts are required as may be necessary to exhibit the causal laws relative to the diverse specific behavioral events.

The positivistic representation of psychological events will seem inadequate to those psychologists, social scientists, and laymen who look in vain among the formal operational concepts for the practical understanding of 'living experience' which is shown

by the novelist and by the discerning man in the street. The best example of this type of dissatisfaction with the natural science approach to psychological phenomena is furnished by the writings of Dilthey and his followers, whose aim is to achieve an understanding of the psychological causes of the cultural events studied by the social sciences. Spranger (**18**) is a leading exponent of this school of thought, and he has attempted to explain the phenomena comprised within the culture sciences (*Geisteswissenschaften*) mainly in terms of the concept of human values or interests. In general, the 'understanding' psychologists seem to rely upon an empathic or intuitive appreciation of the significance of human behavior, in terms of the general forms of its cultural manifestations. The elementalistic concepts which are characteristic of the natural sciences are regarded as meaningless for the purpose of interpreting complex social phenomena. With this last proposition I can entirely agree, but it does not follow that one must therefore introduce a sharp dichotomy between *Naturwissenschaft* and *Geisteswissenschaft,* with respect to describing human behavior scientifically. This dichotomy implies that natural science is to be defined in terms of what Dodge (**4**) has called a propensity for particularizing analysis, a definition which may derive support from the historical behavior of scientists, but which cannot be substantiated by recourse to the general logical precepts of universal science. The fact is, as Lewin has clearly shown, that psychology has not developed concepts which are adequate to the description of the complex phenomena which fall within the general field of social psychology, phenomena which are indicated in part by Lewin's terms "will, needs and personality." But there is as yet no reason to assume that such phenomena cannot be represented conceptually and investigated scientifically at a meaningful level of organization. The 'psychological' insight of the novelist or of the man in the street may be very accurate, in terms of success in the pragmatic regulation of conduct, but such skill is only very imperfectly communicable since the basis of it is largely unknown. By contrast, the skill which is based upon scientific generalization, although often less serviceable in practical affairs than the *private*

skill of exceptional individuals, nevertheless can be more widely utilized and can become a part of the generalized cultural equipment of the race.

Science is grossly misunderstood by the individuals who criticize it because of its abstractness and consequent lack of the concrete richness of 'living experience.' It should be perfectly clear that the scientific account is, as Tolman observes, a *map account*. Each of the various scientific maps of reality represent only those features which are significant for the general regulation of conduct in respect of the segment of reality to which the map relates. If there are aspects of experience which cannot be mapped, then just to that degree are there events which fall without the scope of scientific description. It is better to recognize this limited function of science than to try to force the scientific account to become a poor replica of the whole of reality. Philosophy, too, presents a map account of reality, and in most respects a map far more abstract than those of the sciences, a fact which is not appreciated by all philosophers. But the notion that the philosophic account of experience somehow approaches more closely to "living reality" than the scientific account can readily be dispelled by casual reading in the *Principia Mathematica* of Whitehead and Russell, or in Kant's *Critique of Pure Reason*. Lest this view disturb our philosophers too much, I hasten to add that the philosophic map is a far more comprehensive map than those of the sciences; a world map, perhaps; indeed, a map which may even be colored!

Bibliography

1. Alexander, S. *Space, Time and Deity*. London: Macmillan Co., Ltd., 1927. Vol. 1, pp. 12–13.
2. Balz, A. G. A. The metaphysical infidelities of modern psychology. *Jour. Philos.*, 1936, **33**, 337–351.
3. Boring, E. G. *The Physical Dimensions of Consciousness*. New York: The Century Co., 1933. Pp. xii+251.
4. Dodge, Raymond. Constructive reactionism. Knowledge of reality from a psychological point of view. *Psychol. Rev.*, 1934, **41**, 98–102.
5. English, H. B. The ghostly tradition and the descriptive categories of psychology. *Psychol. Rev.*, 1933, **40**, 498–513.
6. Holt, E. B. Materialism and the criterion of the psychic. *Psychol. Rev.*, 1937, **44**, 33–53.

7. Hull, C. L. Mind, mechanism and adaptive behavior. *Psychol. Rev.*, 1937, **44**, 1–32.

8. Hunter, W. S. The psychological study of behavior. *Psychol. Rev.*, 1932, **29**, 1–24.

9. Johnson, H. M. Some follies of emancipated psychology. *Psychol. Rev.*, 1932, **39**, 293–323.

10. Johnson, H. M. Did Fechner measure 'introspectional' sensations? *Psychol. Rev.*, 1929, **36**, 257–284.

11. Johnson, H. M. Some properties of Fechner's 'intensity of sensation.' *Psychol. Rev.*, 1930, **37**, 113–123.

12. Koffka, K. *Principles of Gestalt Psychology*. New York: Harcourt, Brace and Co., 1935. Pp. xi+720.

13. Köhler, W. *Gestalt Psychology*. New York: Horace Liveright, 1929. Pp. x+403.

14. Lashley, K. S. A behavioristic interpretation of consciousness. *Psychol. Rev.*, 1923, **30**, 237–272, 329–353.

15. Lewin, K. *Principles of Topological Psychology*. New York: McGraw-Hill, 1936. Pp. xv+231.

16. Rosenzweig, S. Schools of psychology: a complementary pattern. *Phil. Sci.*, 1937, **4**, 96–106.

17. Sanborn, H. C. Methodology and psychology. *Phil. Rev.*, 1928, **37**, 15–41.

18. Spranger, E. *Psychologie des Jugendalters*. Leipzig: Quelle and Meyer, 1929. Eleventh Ed.

19. Stevens, S. S. The operational definition of psychological concepts. *Psychol. Rev.*, 1935, **42**, 517–527.

20. Tolman, E. C. *Purposive Behavior in Animals and Man*. New York: The Century Co., 1932. Pp. xiv+463.

VOLUME 50 WHOLE NO. 226
No. 6 1938

Psychological Monographs

EDITED BY
JOHN F. DASHIELL
UNIVERSITY OF NORTH CAROLINA

An Experimental Approach to the Reading of Music

By
KENNETH L. BEAN

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN THE UNIVERSITY OF MICHIGAN
1938

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
THE OHIO STATE UNIVERSITY, COLUMBUS, OHIO

ACKNOWLEDGMENTS

The writer wishes to express his appreciation of the helpful suggestions given to him by the following individuals: Dr. W. B. Pillsbury, Dr. Martha G. Colby, Dr. George Meyer, Dr. John Shepard, Dr. Max Wertheimer, Dr. C. H. Bean, and Professor Wassily Besekirsky. He is also grateful for the services of numerous professional musicians and students who made this research possible by acting as subjects in his experiments.

TABLE OF CONTENTS

AN EXPERIMENTAL APPROACH TO THE READING OF MUSIC

I. The Nature of the Problem

1. *Statement of the Problem.*—The purpose of this investigation is to determine the complexity of the musical pattern that can be perceived at one fixation of the eyes by individuals with different amounts of musical training and experience, and to study the effects of practice with a tachistoscope on the span of perception of these individuals for various kinds of musical material.

In order to make this statement clear, we must define *complexity* as used in this connection. The number of notes included in the pattern is only one of the factors to be considered. The arrangement of these notes on the staff, their relation to one another, and the way they sound when played are extremely important. Thus we may consider optical complexity and musical complexity as two more or less independent aspects of the music reading situation. A group of notes may be optically complete as a gestalt, but musically an incomplete configuration, requiring what music theorists term a resolution. The reverse is equally common. An irregularity in the optical pattern may look bad, but not sound as an irregularity at all. Since most trained musicians transfer the visual impression into the auditory field almost instantly, even before playing the notes, the effects in both sensory fields must be taken into account. How this can be done will be clear to the reader later.

The term *span of perception* seemed to the writer to be the most descriptive expression to indicate the amount of material seen distinctly at a single glance, in other words the number of notes that could be played correctly when exposed for a fifth of a second or less. This number might be constant or variable. Possibly no more than a single note could be read at a time by anyone, regardless of training. These unsettled questions pre-

sented a problem to the experimenter, which was to determine how music is read: whether by single notes or in patterns, and if by the latter method, how the notes are grouped.

2. *Need for Its Investigation.*—For years the writer has observed that many accomplished performers on musical instruments are inaccurate, slow, and stumbling readers of notes. Some of these are professionals of high standing, while others are students believed by their teachers to have exceptional talent. Nearly all of them claim to have put forth persistent effort to improve their sight-reading, with but little success, however. In striking contrast to these individuals, there are some who are not accomplished technically on their instruments, who can read at first sight the notes of anything within the limits of their ability to perform. Examples of both of the above types can often be found among pupils of the same teacher. Poor readers and good readers of all ages are to be found at every stage of development in musical training.

Violinists are, as a rule, faster readers than pianists, for the obvious reason that the former are usually following only one melodic line on one staff, while the latter must read chords on two staves and not infrequently more than one melodic line besides. Singers may be good readers or inferior ones. Many of them are not accomplished readers, even though they have only a melody to sing. Perhaps the most remarkable achievement in music reading is demonstrated by organists, many of whom can play a fairly difficult, unfamiliar composition with but few errors. In order to do this they must respond simultaneously to three continuous, independently varying successions of stimuli. The notes on the top staff must be played by the right hand. Those on the middle staff are for the left hand. The feet operate pedals for notes on the lowest staff. The three parts are often relatively independent as to time relations and direction of motion, even though they fit together according to the rules of composition and counterpoint to form one unified masterpiece. Under such conditions a well-established hierarchy of habits must carry on each mental process for a time without the focus of attention being necessary. The organist cannot be

clearly conscious of three acts at once. Exactly how he can perceive and react to so much material simultaneously is certainly not understood, but the fact remains that he does it.

The orchestra conductor has much more material on his score than he can perceive clearly if it is unfamiliar, but he needs to keep dominant in his thinking only important parts of his score, and leave the rest to his players. Even at that his task, so far as reading is concerned, is often quite difficult. How the notes tend to fall into patterns for him, if they do at all, is unknown.

All of these examples illustrate the fact that the reading of music is infinitely more complex as a mental process than the reading of ordinary printed words. In music we may find a new combination every time we turn a page—a passage that is totally unlike anything we ever played before. It is similar to a strange word that we have to look up in a dictionary. Such innovations are most frequent in contemporary music. Some players reproduce these irregular passages with remarkable accuracy, while others stumble over the most conventional themes with far too many errors, considering their experience.

Successful learning of the skill of efficient reading seems to involve a trick of which neither teacher nor pupil is conscious. Those who are lucky enough to hit upon it make amazing progress, but they seem unable to give much assistance to those less fortunate than they. The nature of the learning process here calls for investigation by psychological experimentation. Teachers are merely guessing when they invent remedies for this difficulty that are based on no good sound theory. In the reading of words the eye moves by jumps along a straight horizontal line. The most efficient type of motion and unfortunate deviations from it are well known to workers in this field. How the eyes should behave when looking at music is a more difficult and as yet unsettled question. Part of the difficulty lies in the frequent necessity of following more than one melodic line at a time, keeping the time relations of the parts as they should be, and at a steady tempo. This requirement is not found in reading words. Transfer of the seen notes to the instrument gives trouble in some cases.

Because of his own difficulty in sight-reading, the writer has interviewed a number of leading music teachers of this country on the subject. Among those who expressed their opinions to him may be mentioned Prof. Leon Sametini, head of the violin department at Chicago Musical College, Prof. Wassily Besekirsky, head of the violin department at the University of Michigan, Mme. Eugenie Wehrmann-Schaffner, nationally known concert pianist and artist teacher at Louisiana State University, and others. The three individuals named above are known to excel in reading as well as in performance. Two other university teachers among those questioned are slow, stumbling readers, but good performers. These teachers all had pupils who were handicapped in reading. They were uncertain as to the cause of their pupils' difficulty and admitted that attempts to remedy the defect had been for the most part unsuccessful. The suggestions given to the pupils differed only slightly from one teacher to another, and may be summed up in the following points:

1. The student should always look ahead of where he is playing. How far ahead he should look is determined largely by the tempo and the time values of the notes. The distance may vary from one beat to two or three measures ahead, if we give extreme cases. The value in following this suggestion lies in anticipation of the adjustment of the position of the hands in preparation to play what is coming next. One teacher has added that this interval of time between seeing the notes and playing them allows the student to digest the music mentally so that it does not take him by surprise.

2. The student should do some sight-reading every day. Any period of time in which this is not done is sure to result in loss of part or all of the improvement previously gained.

3. A large amount of relatively easy material, well within the technical ability of the performer, should be covered in practicing sight-reading. Attempting to read material that is too difficult results in stumbling and looking back to correct many errors. Thus all feeling for the time value of the notes is lost.

4. When sight-reading, one should not stop to go back and correct errors in what has just been played. This quite common practice leads to a habit of looking backward now and then instead of looking steadily forward for what is to come next, and naturally stumbling results. It is not easy for some careful performers to drive on through a piece in time when they know they are playing wrong notes occasionally, but they must force themselves to do this even though inhibitions set up because of making one error may bring about more errors. If they do not drive ahead in spite of mistakes, they will always read slowly. Of course the student must refrain from applying this principle to practice on compositions he wishes to perfect. Such disregard of errors is helpful only in increasing speed of reading.

Careful observation of the reading of skilled musicians reveals the fact that much guess-work is involved. The player sees a few cues and fills in what seems to be appropriate to complete the pattern. The effect usually sounds correct, but the notes played often fail to conform exactly to what is written. The player sees three notes in a chord that suggest to him the dominant seventh for example. Without taking time to see what the other notes of the combination are, he fills in the rest of this harmony in the position or inversion which seems to him to follow logically after what he has just played. He is likely to guess right or nearly right as to what notes belong there. Irregularities that are inconspicuous in scale or chord passages are likely to be overlooked and played as if part of a regular, conventional pattern, however.

When asked how they acquired skill in their own sight-reading, these leading teachers could think of no additional suggestions besides the above, but merely admitted that they did not know exactly how they read so fast or why some of their pupils could not do the same. The writer then suggested what was at the time his hypothesis that perhaps they saw notes in groups and not as individual notes. To this three teachers responded vaguely to the effect that sometimes they probably did see groups as wholes and not as individual notes, but that they were seldom conscious of how they did it, and had no idea of the extent of the groupings. The other teachers did not know how they read.

Teachers of public school music are now making use of what are called *flash cards* in teaching sight-singing. They hold before the class a large card with four notes or more on it, then take it away and ask the class to sing the melody. The exposure time is variable, and the technique crude, but it seems to be a step in the right direction. As no accurate measure of the results of this teaching method is now available, any detailed review of literature on sight-singing methods would have no bearing on the problem of this research, and will therefore be omitted. The fact remains that public school music teachers have no proof that the children learn to perceive as wholes the groups of notes that

are pointed out to them. An accurate measure of progress is needed, which must be the result of extensive experimentation.

3. *Its Practical Value.*—In any kind of professional work in music, whether playing or teaching, a large amount of material must be played which the performer has no time to memorize completely or even partially. Good technical background is not the only prerequisite for this sort of playing. A certain level of efficiency in reading is essential if much extra work is to be avoided. Those who play an instrument for pleasure should not be confined to what they can reproduce from memory for their enjoyment. They should be able to enjoy reading new music as they delight in a new novel. This is impossible for many, however, because their slow, stumbling reading gives rise to a feeling of fatigue and unpleasantness, and in extreme cases an emotional complex, causing complete inhibition at the sight of printed music. Often a slight deficiency is made much worse by an over-severe teacher who "explodes" at every wrong note the pupil plays. The resulting mental attitude takes all the pleasure away from reading the printed page, and causes a careful note-by-note analysis of all new music, which is necessarily slow. Frequent backward looking to see if all is going well results from a fear of making mistakes and provoking the teacher's disapproval.

The practical aim, then, of this investigation is to help the teacher to find the individual difficulty of each pupil in music reading, and to correct the defect. This problem is essentially one in applied psychology, though its theoretical aspects are to be brought out also. We do not claim that the secret of efficiency in the skill under consideration lies entirely in the application of any one principle of psychology. Complex skills must go through many stages before perfection is finally attained. The child must learn what to do in response to every note on the staff when he is a beginner. He must automatize this response so that he does not have to think about it. He must attain some speed in manipulating his instrument. Then comes a final stage in the learning process having something to do with rapid perception

of notes and quick, well coördinated response to them. This step is the one least understood, and the one in which the writer thought the gestalt principle might play an important part. Out of such an idea the writer got the basis for his experiments to find out what we do with music when we read it.

II. Discussion of Literature Related to This Study

1. *Studies of Music Reading.*—For nearly twenty years music teachers have pointed out certain note groups to their pupils. They have asked the pupils to learn how these groups look, but they had no measure of the pupil's ability to respond to these groups as wholes. In fact they had no notion as to what patterns of notes they themselves read ordinarily. Various methods of teaching music reading by sight singing have been attempted. The use of the conventional fixed *do* system was followed by the American movable *do* with syllables for chromatic tones as well as those of the diatonic scale. This was complicated, and whether it was any better than the older system may be questioned. Numbers have been substituted for syllables in some schools. These systems have brought results, but work with them has been based upon opinion and not upon scientific experiments. Therefore detailed consideration of them will be omitted here. The study of harmony and counterpoint acquaints the student with patterns in music in greater detail than merely playing an instrument. The specific effect of such study upon reading is not understood, and has never been measured. The crude methods of note group reading on flash cards, recently used by some teachers, measure nothing with any accuracy as to progress, and the selection of material for this purpose at present is not based on experimental study. However, a more careful use of flash cards may prove to be a step in the right direction.

Possibly the clearest advice given by any musician on reading is to be found in the following quotation from Joseph Hofman (6): "For improving the facility in sight-reading, you must do much reading or playing at sight as fast as possible even though at first some slight inaccuracies may creep in. By quick readings you develop that faculty of the eye which is meant by 'grasp,' and this in time facilitates your reading of details. . . . A large part of sight-reading consists of surmising, as you will find upon analyzing your book reading."

This statement suggests pattern reading vaguely, but gives no clear explanation of exactly how it is done.

Scientific research on music readng has been done by Jacobsen (7) and later by Ortmann (9). Jacobsen photographed eye movements with a horizontally moving film for one eye and a vertically moving film for the other, in order to be able to pick up motion of the eyes in any possible direction. He used ten adult subjects. The music was played on a small reed organ. Only one of his subjects had ever played organ before. Errors were recorded. If a whole chord was played wrong, this was counted as only one error. Fixations were found to be short in good readers, and few regressions occurred. If the reading speed exceeded the playing speed, some mistakes resulted. Errors crept in also when too large a pattern was attempted, because the group was not grasped well.

If the tempo of the music was slow, one fixation occurred for every one or two notes, as often one as two for good readers, for one-part material. If the

melody was faster, three notes on an average were taken in with one fixation by good readers. Poor readers made one fixation per note, or even two on long notes, with some regressions to make sure of the key signature. Often a fixation was not right on a note, but near it. Fixations for all subjects were more numerous at a normal rate than when speeded up by a metronome. The number of fixations for a simple melody ranged from 30 to 62 in the group. There was one fast reader who made up for his excessive number of fixations by making them very short. Average length of a fixation varied individually from .32 second to .76 second. One subject had extremely long fixations, but few of them. He was a slow reader, however.

Two-part material worked out in a similar fashion, but the errors were all in the lower part, which, naturally, was taken to be of secondary importance compared with the top melody. At least the subjects seemed to give more careful attention to the upper part. An average of one fixation for each pair of tones was made, and 30 to 48 pauses were necessary to read a two-part selection.

Reading of three-part material required many movements horizontally, vertically, and diagonally from the one staff to the other. The reader is referred to Jacobsen's interesting diagrams to see exactly how this took place. It was evident that the performer had to remember one part while playing the other, and look ahead to anticipate what was to come next. Jacobsen found that little guessing was done. Everything was perceived. Any irregularity in a pattern, such as varied rhythms, accidentals, etc., caused an increase in the number of pauses. If errors were made, the same effect was observed on the film. At a normal tempo, the players averaged 67 pauses for a three-part selection. The same at a faster tempo gave an average of 47 fixations. The good reader, according to Jacobsen, looks considerably ahead of his playing, but the poor reader is not much ahead. The experimenter did not give exact data on how far ahead each looked, but such facts might be interesting.

Jacobsen found that the less experienced readers must take more time to get both hands successfully than to grasp notes for only one hand on one staff, but the mature reader can maintain the same speed for reading on two staves. The number of regressive movements back to groups that were probably forgotten increased with the three-part material.

Four-part music illustrated the same principles, except that more notes had to be grasped per fixation in order to play it up to time. Good readers grasped as many as six notes at one pause, according to Jacobsen. The average span of the subjects was much less than these exceptional instances, however, varying individually from 1.95 to 2.97 notes per fixation. Errors did not increase in proportion to the number of parts added, but most of them did occur in the bass for the four-part material as before. In this case the fixations on the two staves were about equal in length and number. Movements were in all directions as before.

Good readers were found to read chords from the top down, not from the bottom up, as a rule. Few exceptions to this were found. Poor readers look at one note at a time until they have built up the entire chord. They are likely to begin at the bottom as often as at the top.

Jacobsen found that his subjects who had relatively little training did not make many more fixations than the highly trained performers. Some of the former made fewer pauses than the latter. All of his subjects made errors on this simple material. The fastest readers were the most accurate, however,

and they differed only in the fact that some of them made many short fixations, and others made fewer, longer ones. In every good reader, retention proved to be a large factor.

Ortmann (**9**) made a study of the reading of 45 chords. He exposed each 2/5 second to each of his subjects, who responded by writing the notes. Two to six notes were included in each chord, which was not necessarily in harmony, but might be any consonant or dissonant combination. The observer seldom failed to get the correct number of notes, but he was often uncertain as to their location. More errors were made if the group of notes was spread over an octave or wider than if it covered a fifth or less, regardless of the number of notes. There was a tendency to underestimate the vertical distance between the notes. Horizontal dimensions were sometimes involved when two notes were on adjacent scale steps so that they could not be written under each other. Interpretations of the results stressed the importance of the visual pattern. Subjects did not think of the tones or the names of the notes. Some looked for octaves and fifths, around which the other notes grouped themselves. Several possible correct ways of writing the same chord occurred, but ordinarily the subject reproduced the design as he saw it. Some saw most clearly the middle part of most chords, or the part where notes were closely grouped. Introspections revealed that 96 per cent of the subjects read from the bottom up except for close groupings, in which direction was unnoticed.

Ortmann (**9**) suggested that chord unit reading is analogous to word reading as against mere letter reading. If a "good chord" was shown, the results were better than for an unfamiliar grouping. Notes were more frequently misread by one step than by two or more. Mental set had an effect. An unexpected chord was missed usually, while if a familiar one came along, it was read easily. About 15 per cent of the observers did not think of the clef used, the names of the notes, or the chord, but only the visual pattern. Ortmann says, however, that more was involved than mere visual grasp. He admits that he covered only a part of the process in his study. Interesting individual differences appeared in the results. Some good readers and some poor readers served as subjects. The amount of training varied considerably also.

2. *Criticism of These Studies.*—In the opinion of the writer Jacobsen and Ortmann have both given some valuable evidence in favor of the view that good readers of music read patterns of notes, while inefficient ones read individual notes. They have not gone far enough, however, to show how various kinds of music may be grouped by different readers. Their material was limited. The individual differences among music readers at all stages of advancement is an extensive field for study by no means covered by these two investigators, who used a limited number of subjects.

After examining the record of the errors made by Jacobsen's subjects on his relatively simple selections, the writer has concluded that Jacobsen did not have any excellent readers in his group. There are numerous musicians everywhere who could have exhibited better performance, and some record of how they achieve it would be interesting. A greater number of subjects should participate in such an experiment in order to be safe in generalizing about good readers or poor ones.

Jacobsen's selections demonstrated effectively what is done in following one, two, three, and four simple parts, but few rhythmic variations entered in. The material was not extensive, nor was it varied in nature. The questions of how complex rhythmic relationships between the parts may be perceived or what

the reader does with a seven or eight-note chord are not yet touched upon. The study of the effect of practice in reading conventional and ultra modern types of music upon the eye movements made with these kinds of material would throw some light upon the part played by past experience. These possibilities have been pointed out in order to show that Jacobsen has described only a small part of the behavior of a few music readers, and that his conclusions may be valid only with the kind of music that he used.

The writer objects to Jacobsen's use of a small reed organ on which his subjects played the selections, because all but one subject were unfamiliar with the organ, and could have done better on the piano, which they had played before and to which they were accustomed. How much the unfamiliar instrument distracted their attention from the reading would be difficult to determine, but it might easily have interfered with good results.

Another criticism that might be made is that the type of camera Jacobsen used, with a vertically moving film for one eye and a horizontally moving one for the other, is subject to certain inaccuracies. Without doubt it is complicated, and some claim that such a camera is unsatisfactory. If the eyes of the subject always moved together and fixated the same point, possibly other sources of error would be too small to affect the general result, but the writer questions whether Jacobsen was safe in assuming that he had located the fixations correctly in spite of the defective coördination so frequently found in the two eyes of many individuals.

A consideration of Ortmann's work shows that his material was limited in scope. His patterns were in many cases mere chance combinations, unlikely in music. He did not have his subjects play the patterns on an instrument, which is the usual response in reading music, but he had them write the notes. This required more time and effort, and was not the same mental process as the typical reading and playing situation calls for. In order to write a chord, a vivid memory of it must somehow be retained for a much longer time than would be required to play it. Also one may observe from the results that many subjects did not think of tones, letter-names, or clef signs, but only of visual patterns. In such a case were these patterns music or were they mere visual forms? One can answer this question by stating that unless the visual forms meant tones to be heard or played by the subject, the response was not that of reading music as we usually think of it. Possibly this difference in the response accounts for the disagreement of the two investigators as to whether chords are usually read up or down. Chords should be thought of in terms of the clef sign used, the key signature, and the use of the two hands (or one in some cases) to play them, etc. The visual pattern is important, but it is not the whole process.

A question might arise as to why Ortmann exposed his chords ⅖ of a second. This is twice as long as one fixation of the eyes. One could not discover whether the subject took in the form instantly as a whole or read it by parts, since either method was possible with this long exposure. A shorter time for seeing the form would have given a more accurate measure of the span of perception.

In spite of the disadvantages of Ortmann's method, however, his observations concerning the nature of the errors are interesting and valuable. He suggested the idea that chord reading is analogous to word reading, and presented some evidence to indicate that this is true, but he gave little in the way of an explanation of how we learn to group notes into certain patterns. His material was not extensive enough to furnish adequate description of the chord reading process, which needs further study in a more typical situation.

III. Apparatus, Materials, and Procedure

1. *The Twin Tachistoscope.*—Since the subjects in this experiment were expected to respond to musical patterns by playing them in most instances, a tachistoscope that could be mounted on a piano was necessary. The use of a tachistoscope of the usual sort placed beside the piano would require the subject to turn his head far out of the natural position for reading music at the keyboard, therefore any such plan was abandoned. The experimenter felt that the conditions for the work should resemble the ordinary music-reading situation as closely as possible. Therefore he designed apparatus that would make these conditions conveniently possible.

Also the experimenter saw the value of testing more than one method of responding to these musical stimuli, because the subject might in some cases experience difficulty and confusion in attempting to transfer what he saw to the keyboard, even though he perceived accurately what was there. How much of the subject's difficulty was in reading and how much of it was in the mere mechanics of playing would not be easy to establish without some kind of a check on this point. The experimenter believed that a comparison technique would be useful for this purpose. Two stimulus cards might be presented in succession, with equal times of exposure and a constant interval between. The task would be to tell whether the two patterns were alike or different, and, if possible, what the difference was, if they were unlike. If the subject could notice a small difference every time one occurred, he was seeing the whole (the configuration), and if he could locate the difference exactly, he was getting the details of the pattern as well as its general form.

This type of comparison experiment could be done with rather elaborate apparatus of a sort now in use in some laboratories, except for the difficulty of mounting such a bulky instrument upon the piano for the other method of response. Conditions should be as much alike as possible in the two kinds of experiments, therefore the same machine ought to be used throughout. The Twin Tachistoscope has the advantage of being mountable on any piano, simple to construct, easy to operate rapidly, seldom in need of repair, and accurate enough to give reliable scientific data.

The instrument, which is illustrated in Fig. 1, is entirely of wood except for the wheels, braces, and smaller parts. The screen is made of 3/8-inch five-ply wood, for durability with a minimum of weight. The blocks, which rest at the ends of the keyboard, are adjustable by sliding them along the board fastened to the lower edge of the screen. The metal braces, padded at the ends that clamp to the piano, are adjustable to pianos of different height. These give additional support to the instrument so that it will not fall forward upon the subject. The two windows in the screen, through which the stimuli are seen, are 2½ inches square. The screen is 8 inches above the keyboard to allow the hands ample freedom for playing. The discs are 20 inches in diameter, and are constructed of heavy galvanized iron painted black on the back side to reduce glare. On the front side, toward the subject, the discs are painted gray to reduce the amount of change in intensity of the light when the card is exposed. If the discs were black on this side, the contrast involved in each flash would be greater, a phenomenon which tires the eyes of some individuals. Bicycle hubs reduce friction to a minimum, and yet hold the discs rigidly in position so that they do not sway when in motion. Mounted on each disc is a fixed weight of lead in a galvanized iron pocket and a movable counterpoise weight that can be tightly fastened by a screw at any desired point along a bar on the radius of the disc. This weight, which is of the same material as the

Front View of Twin Tachistoscope mounted on a piano.

Subject reads card flashed in left window, then plays the notes.

Experimenter at left has just pulled the trigger.

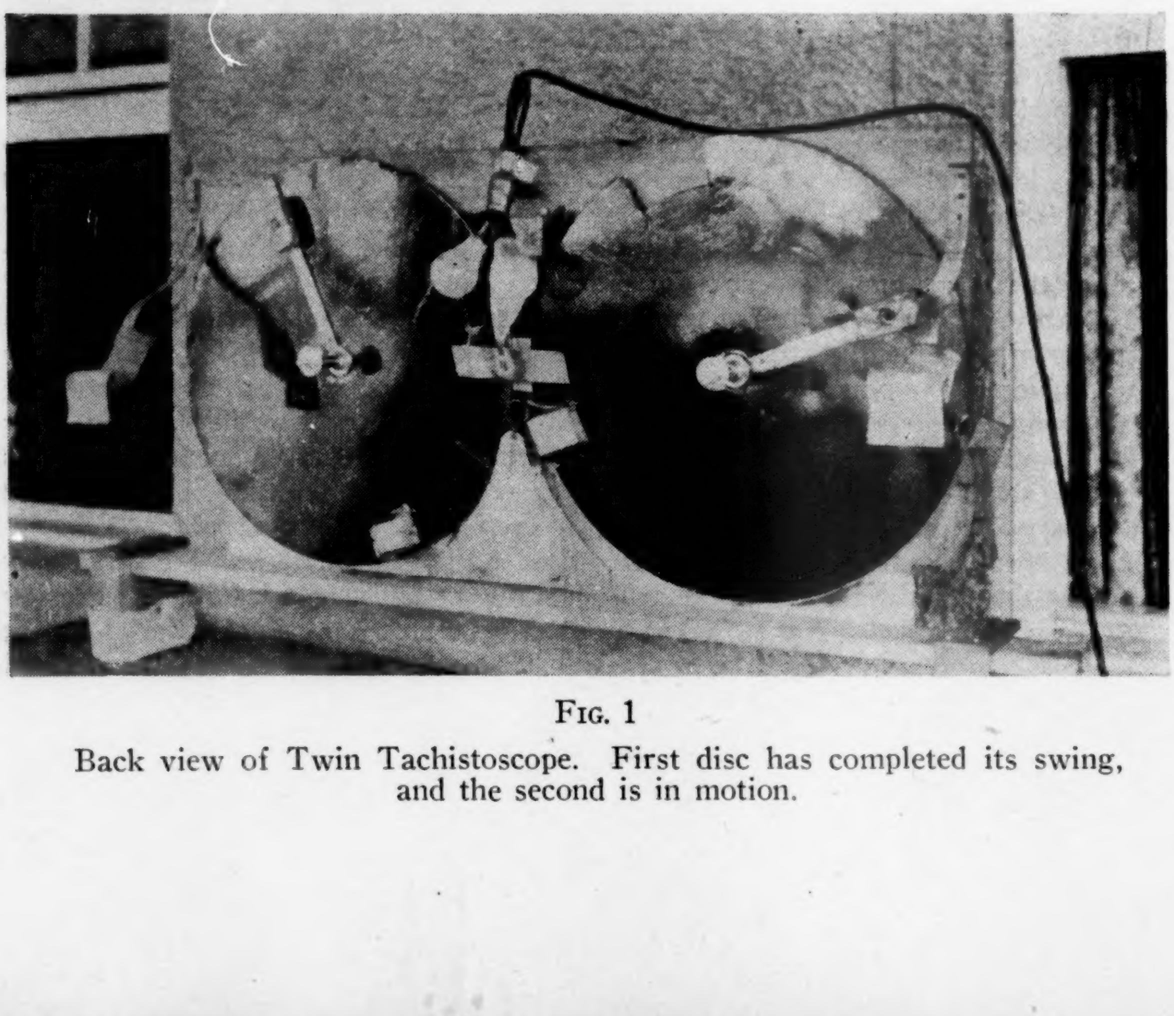


FIG. 1

Back view of Twin Tachistoscope. First disc has completed its swing, and the second is in motion.

other, but slightly lighter, regulates the speed of rotation. The gap in each disc is so cut as to expose the card during the fastest motion of the wheel as it swings. This means that a diameter drawn perpendicular to the one on which the two weights are located passes through the exact middle of the gap. Since the wheel works on the principle of the counterpoise pendulum, the fastest motion occurs at about the time the gap passes the window, making the opening and closing of the exposure as quick and inconspicuous as possible. Any differences in weight or form of parts of the two discs that may have escaped notice in spite of great care in construction were too slight to cause any inaccuracy large enough to be observed.

At the extreme right of the back view of the apparatus may be seen the trigger which releases the fixed weight of the first disc, allowing the weight to fall without a push from any part of the trigger. After it has swung to the other side, the weight is caught by the latch on top of the small triangular board between the discs. No backward swing is possible until the experimenter pulls the latch into its original position. If the second disc is to be used, it is set before the first one is released, in such a position as to fit the left prong of the latch into a notch in the disc. Thus set, the weights are in position for a swing opposite in direction, but exactly equal in extent to that of the first disc. When the fixed weight of the first disc strikes the latch, its left prong is jerked out of the notch in the second disc, the fixed weight of which falls and is caught near the end of its swing by two springs fastened to the latch board. This disc remains there until set for another swing. All parts operate silently, since rubber padding was used at all points of contact of the latch, springs, or trigger with any parts of the discs.

The regulation of exposure time was accomplished by marking points along the bar at which the sliding weight could be set for certain desired speeds. The determination of these points was made with a Hipp Chronoscope, for which electrical contact was made and broken by a device constructed for the purpose. The contact system consisted of a metal bar fixed to the edge of a window in the screen and a flexible spring that would touch it unless kept a very short distance from it by pressure against the disc. When the disc was in motion, the spring, which was small, made a negligible amount of friction against it, but enough to keep it 1/16 inch from the bar, approximately. When the gap reached the window, the spring was instantly released against the bar, making contact, which was later broken by pressure of the other edge of the gap against the spring, quickly pushing it off of the bar. For the purpose of this experiment, an exposure of slightly less than a fifth of a second was desired, since this has been considered the best time for tachistoscope work. During such a brief exposure, the eyes cannot shift from one part of the stimulus card to another, but must take in everything at one fixation. Thus a measurable unit of perception can be obtained. The range of times possible with the Twin Tachistoscope was found to be from about 2/3 second to 1/20 of a second, but the most desirable exposure turned out to be .187 second as an average of 20 trials. The total range of variation of this time did not exceed .005 second for these 20 measurements. Much work was required in reducing friction and adjusting nuts on the hubs before this degree of reliability was finally reached. More tinkering was necessary before the second disc was at last synchronized with the first so that its exposures averaged .185 second in 20 trials with equally small range of variation. Of course crude synchronization was achieved first by allowing both discs to swing freely many times, adjusting the sliding weight of the second until it went with the first for several swings without perceptible deviation. As soon as the speeds seemed the

same for some time, the contact device was shifted to the other window for a test, with the result that, after further small adjustment, reasonable accuracy and dependability were attained.

The card support is of galvanized iron, and can be swung out slightly to facilitate card mounting. Considerable care must be taken to avoid setting up a vibration of the support when releasing it with the hand, but with practice this can be done easily. The support is long enough to hold a card with two stimuli, one before each window, but it may be used for a single stimulus card to appear in one window only. The card is held at a slant that is about parallel to a sheet of music at the rack of the piano, and is about a half inch behind the window at the bottom.

Illumination is furnished by two 60-watt bulbs in a double socket, projecting from the screen at an angle such that there is no glare on the stimulus card at any point. The lights are not visible to the subject when the exposure is being made or at any other time. Care is exercised to have the room dimly lighted with no glare anywhere that will shine in the subject's eyes. Intense light on the front of the screen is avoided, since under such a condition the subject's eyes would tend to follow the small irregularities in the surface of the discs as they turn. Such irregularities are inconspicuous in dim light, and steady fixation is undisturbed. Many lighting schemes were tried before a successful one was found. The importance of eliminating all distracting stimuli and glares, even though slight in intensity, was revealed in preliminary experimentation with the Twin Tachistoscope before any results for actual use were obtained. No set of conditions will please everybody. The 60-watt bulbs gave too much illumination for some individuals, but reduction of the intensity caused others to complain that the cards were not sufficiently lighted. In general subjects quickly became adapted to any feature of the apparatus that bothered them during the first few trials.

A problem arose, in connection with the use of the machine, as to how a satisfactory fixation point for the eyes might be located. Workers with the tachistoscope who are careful about this matter have the subject fixate a point in the middle of the area to be covered by the material to be seen. This seems to be a logical procedure, and works out experimentally for the best if the material consists of a word or a row of dots. Unfortunately, however, the musical patterns for this research were irregular in form in many cases, and almost never on a straight horizontal line like a word. Therefore the arrangement of a group of notes on a card in such a way that a fixed point of gaze would always fall in the middle of the group was next to impossible. Where the middle would be was in some cases debatable. Yet the experimenter knew that wrong fixations would cause a part of the pattern to lie outside of the area of clearest vision, and thus lead to errors. Therefore he placed a square of smooth, clear celophane over one window, with a black ink dot in the middle of the square. The dot was very small, but clearly visible before exposure was made. Applying a wet, sticky mixture to the edges, the experimenter was able to shift the celophane easily and have it stay where he put it. As a part of the preliminary experimentation, he then tried to find the best fixation point for himself and later for two other individuals on several musical patterns. The results were so confusing and conflicting that the establishment of a best point of fixation for each pattern seemed to difficult to be worth the time and trouble necessary to do it. Therefore this phase of the problem was set aside in order to find out more important things about music reading first. Some means of control of fixation that will be adequate for reading music on cards will have to be worked out as a separate experiment in the future.

In reading notes from the page, the performer sees in the periphery of his field of vision whether what is to come next goes up or down. Probably he uses this vague cue to anticipate what he is to do next, at least to the extent of knowing where he must look for the next group of notes. He is playing some notes below the staff, for example, which at the time fall on the fovea, and are seen clearly. While doing this, however, he perceives vaguely, out in the periphery of the field, that the next measure jumps to the top of the staff and runs up above it. He anticipates this change, and shifts his eyes to the right point to bring out this next group clearly.

Reading in the Twin Tachistoscope, the performer has no opportunity to anticipate the direction of motion of the music before he actually fixates it. This leads to many wrong fixations, and reports such as, " I didn't see that at all," unless verbal directions are given as to about where to look. Therefore if the music about to be shown changed from below the staff on the preceding card to above the staff on the one to be seen next, the subject had to be told to expect notes above the staff. Even with the greatest of care in cutting and pasting music on the cards so that the staff would fall in the same location in the window, and with directions where to look when this seemed necessary, some errors were undoubtedly caused by wrong fixations, and not by inability to perceive as a meaningful pattern what was presented. These cases constituted a minority of the errors, however, and were identified introspectively as wrong fixations by the subject as far as it was possible for him to do so. In some instances the subject was sure that he had looked too high or too low, and another trial on the card was given him. If he was not sure this caused him to fail, no second chance was allowed. Unfortunately this uncontrolled factor in the use of the apparatus could not be avoided, but the writer believes that it affected the results very little. The Twin Tachistoscope has proved itself of value, in spite of a few imperfections, for accurate measurement of music-reading efficiency and progress in both speed and accuracy. It has made clear many individual difficulties not understood by music teachers, and has helped to correct most of these faults.

2. *Selection of Material.*—The music to be read was mounted on cards 3 inches long and 2 inches wide, care being taken to have the staff, or two staves in some cases, in identical position on every card. Printed music was cut up into fragments, which were put on the cards smoothly with Duco Household Cement. In a few instances, printed music could not be found which would illustrate a desired type of pattern well. Therefore manuscript patterns were made with a pen as nearly like printed music in form, size, and every other respect as possible. The notes were of the average size found in most printed sheet music. Variations in the size of the staff were too small to make any difference.

The material was divided into three types of music: (1) melodic, (2) polyphonic, and (3) harmonic. Under each heading came groups of cards of increasing difficulty from the simplest possible to the most complex. Sixteen series of cards

were prepared for practice, and an extra series, half of which was testing and half for trial previous to practice. The total was 680 cards. Each series contained 40, with the exception of C and D, which had 50 in each. Series A, B, C, D, and E were made up of melodic passages only, increasing in difficulty from one series to the next. Polyphonic music (having two or three relatively independent melodic lines) furnished the material for series F, G, and H. Part of this music was on one staff for one hand and the rest of it was on two staves for both hands. The latter was far more difficult to read in a short time. Here again the complexity increased as in the melodic type. Chords of two, three, four, five, and six notes (or even more in a few cases) were selected for series I, J, K, and L. Some of these were for one hand, and some for both. They included triads in their various inversions, seventh and ninth chords, and altered chords, etc. Unclassified material of unusual sorts, rhythms, arpeggios, and queer modern melodies fell in series M, while N, O, and P were double exposures for comparison as explained before. These last three covered the three types of musical patterns of the previous series, with fewer examples of each, however.

Some of the melodic patterns were composed by the writer to illustrate the relative difficulty of perceiving certain intervals between notes, and to show the effect of distribution of the notes on the staff upon the ease with which they may be perceived. Other brief portions of melodies were selected from the Bach Sonatas for violin unaccompanied. These passages were chosen because of their good melodic character and their freedom from accompanying harmony. Melodies of a more modern style were selected from the works of John Alden Carpenter and Samuel Gardner, both contemporary American composers. Bach again came into our polyphonic series, naturally. Chords were selected wherever they could be found arranged as desired.

The comparison stimuli duplicated in type, but not in exact notes, each kind of pattern used on the single cards. The checking of one technique against the other which this made possible proved to be of great advantage in understanding the results. The selection of stimuli was not a mere random choice. Each

pattern of notes used illustrated a point in regard to the difficulty of integrating a certain type of configuration. Professor Max Wertheimer, who is accomplished in music as well as in his own field, psychology, gave the writer a number of suggestions as to what constitutes complexity in a group of notes. He pointed out that one must distinguish between optical completeness and musical completeness of a pattern. A configuration of notes may look complete and sound as if something is missing that should bring it to a satisfactory conclusion. Likewise a pattern may be all there as far as our auditory impression is concerned, but appear visually incomplete. Optical irregularities in an otherwise symmetrical ascending and descending passage may have a profound effect upon its perception as a whole, according to Wertheimer. Numerous illustrations of the phenomena described above were discovered and added to the material.

Dr. Martha G. Colby, also accomplished in music as well as psychology, suggested the use of patterns from contemporary music that would not be conventional in form so that the part played by past experience would be equal for all subjects, since the patterns would be new to them. Some such material was included in series L and M. She also pointed out the importance of control of the tempo at which the music was played. In preliminary experimentation, the writer found rigid control of the tempo by means of a metronome undesirable for the kind of practice to be done. Attention to a metronome or to keeping a given tempo too exactly distracted the subjects, seriously hampering them in their attempts to get as large a group of notes as possible at a glance. Therefore control of tempo was confined to the simple instruction: "Play it at a convenient, moderate tempo." This direction was carried out with fair success.

Preliminary trials also revealed that the cards were read as isolated units, and not as parts of a logically continuous composition. Reversal or rearrangement of the order of a logical passage read by single cards did not confuse the subjects or materially reduce their efficiency in reading, strange as this may seem. That fact was favorable for the procedure used, since on the difficult cards the subjects made errors mostly at the end,

making the material sound illogical to them anyhow. Therefore succession did not operate as it does in continuous reading, as an aid in the anticipation of what is to come next. Hence logical order was not followed strictly in all series, though in some the material was in succession as in the composition chosen.

The piano technique required to play the material selected was limited, so that all subjects who play with a fair degree of efficiency should find the response to these stimuli no more difficult than pronouncing a familiar word correctly. The motor response was further facilitated in that the subject was allowed to play the card again if his finger struck a key not intended. This seemed only fair, since the best performers make such errors occasionally, and correct reading rather than playing was to be considered.

Fig. 2 shows examples of the music which the subjects read. The material as shown here is crowded together to save space, and is not distributed as it is on the cards. The illustrations are in smaller notes than the average sized printed ones on the cards. These examples will be referred to again in discussion of the results.

Difficulty was encountered in preparing material because of the presence of distracting stimuli such as fingerings, signatures, accent marks, etc., on nearly all music. Better results might be obtained with no markings, just notes and staff, on the cards. On the other hand, these markings are a part of the pattern in ordinary music reading just as punctuation belongs on a page of printed prose or poetry. Leaving them out would make the situation unnatural. Many markings of expression and fingerings must be disregarded in a first reading of a composition, because attention is on the notes and their time values if the music is at all difficult. Therefore some markings were allowed in the material, but the subjects were not required to observe them. Since holding a signature in mind is essential to good reading, signatures of several sharps and flats were included. The experimenter gave these to the subject verbally before presenting the first card to him, and made note of any failure to observe the signatures throughout.

FIG. 2
Samples of music selected for the cards.

FIG. 2—CONTINUED

FIG. 2—CONTINUED

Illustrations of the same chord or polyphonic figure played with one hand and later with both hands were scattered through some series for comparison of the difficulty of reading the same music written two ways. Comparison of the success in reading treble for both hands and treble for the right hand with bass for the left was also made with the same patterns, not presented successively, however.

Though many principles were to be illustrated in the selection of stimuli, the writer believed that he had collected enough cards to prove something about each point. More thorough investigation of some of the details would be desirable, however.

3. *The Diagnostic Experiment.*—In order to make a survey of music readers of all degrees of efficiency and with varied background of experience, the experimenter selected 50 individuals to be tested. Of this number 28 were men, 19 were women, and three were children (two boys 10 years old and a girl aged 11 years). The familiarity of these people with printed music ranged all the way from no knowledge of it at all to the keen comprehension of professionals of long years of experience. Many degrees of efficiency in reading music were found at all stages in musical education. As a rule at least an hour was spent in measurement on each person, but in some cases two hours or more were required to diagnose a subject's difficulties and determine to what degree he responded to patterns of notes if at all. In this time the subject read as many cards from each type of material as he could, while the experimenter made a careful record of every error, to be checked later.

A scheme similar to shorthand for recording as nearly as possible the exact location of every error was invented by the experimenter, who found its use a time saver after some practice with it. Since he has absolute pitch and experience in analysis of music by ear, the experimenter found that looking at the keyboard while the subject played was not always necessary. Furthermore the writer had the advantage of knowing thoroughly the music he had selected, while his subjects had never heard any of it. Hence all errors, even playing in the wrong key, registered immediately in the experimenter's consciousness. If the

subject missed the first note on a card, the last, the last two, the highest, the lowest, a sharp, a flat, or the time values, a separate marking in the appropriate place on a prepared data sheet kept a permanent record of the error. If more than two errors were made, only the number of them and whether they were in pitch of the notes or in their time values were recorded. This procedure took little time, even though a large variety of markings was employed. Introspections about particular cards or persistent difficulties were frequently recorded on the back of the data sheets. Later, by looking at each card read, the experimenter could tell from the record of the subject what types of errors were most persistent. He could also calculate the average span of perception for each kind of material read.

When a person was selected for diagnosis, he was given some preliminary trials with the apparatus before any records were taken. During this time he learned where to look for the stimulus and became accustomed to the flash of the card and motion of the disc. He was told the signature, time, and whether the right hand, left hand, or both were to be used. Then the experimenter said, " Ready—go," allowing about a half second between the two signals and releasing the trigger precisely upon saying " go." Not later than a second after the flash, the subject was supposed to play what he saw and to guess if not sure of all of it. Pondering over any pattern or reconstructing it from the afterimage was discouraged, since this method would not be possible in straightforward reading. As soon as the exposure ended, the card was quickly removed, while a pencil was taken up to write the record for that trial. After writing the result, the experimenter swung the weight back for another trial and put in a new card. Immediately a pair of signals was given as before. Steady rhythmic action was kept up to maintain regularity, except when it was interrupted to record introspections. The subject was not told whether any of his responses were right or wrong until an entire series was completed, at which time he was allowed to examine his record.

Conditions were kept reasonably quiet during experimentation. The first part of the work was done at the University of Michigan

in a laboratory that was far removed from almost all noises that could serve as distractions. Later the experiment was set up in a 90 per cent sound proof studio in the Louisiana State University School of Music. Still later it had to be moved to the psychology department of the same institution. Similar conditions prevailed throughout. Good pianos were obtained in all three places and put in excellent condition before work was begun.

Subjects who did not have enough knowledge of music to play the piano were asked to state what line or space the notes were located on, or to draw the pattern on a prepared staff. They were given the comparison cards as the others were, but could not name the notes that were changed. They merely located the changes. The purpose of using nine of these untrained people was to determine whether they saw the notes in patterns, and whether they could reproduce what were for them nonsense visual figures or notice differences in the details of similar ones. All other subjects could at least play the simpler cards.

4. *The Practice Experiment.*—The purpose of this portion of the work was to study the general effects of practice (not on any specific material), on the ability to organize patterns on the stimulus cards presented. Preliminary research indicated that some kind of Gestalt is involved in reading music. Therefore the experimenter felt safe in proceeding on the assumption that a Gestalt interpretation of music reading is possible and might prove useful.

The total number of subjects who practiced was 24, of whom 14 were women, 8 were men, and 2 were high school girls, one aged 13 years and the other 14. With the exception of 3 the ages ranged from the late teens to 30 years. One middle-aged woman was in the group. Nearly all were advanced technically on the piano, and could play material of fourth grade difficulty at least if not more advanced repertoire. Two of this number were not technically efficient enough to make progress beyond a certain point possible, because their response to notes was not sufficiently automatized. These two only practiced 10 hours each. Nine practiced 20 hours, and the remaining 13 did 30 hours of work each. Unfortunately not all of the subjects could

be obtained for as much as 30 hours of work. One in the latter group worked 3 hours beyond the 30 out of her own interest to see if further progress would be possible. The greater number of females could not easily be avoided since comparatively few men study piano in college.

Selection of the individuals for the experiment depended upon their present ability to play and their past training. Ordinarily four years or more of previous piano study proved to be essential, so that the motor response was never too difficult. Some knowledge was obtained about each person's training, efficiency in playing, and accuracy and speed in music reading before the experiment started. Piano teachers furnished this information in most instances, but in addition the experimenter asked the subject to read at the piano, before the tachistoscope was mounted, two selections, the first slow in tempo and the second fast. Piano interludes in " Symphonie Espagnole " by Lalo for violin and piano were used for this purpose, since almost no pianists have ever played it unless they are expert accompanists. A subjective evaluation of the reading of each person was then entered on his data sheet. He was rated in a manner similar to the grading system used in most universities: A, B, C, D, and E with plus or minus signs in a few cases.

Following this test he was given a brief explanation of the apparatus and purpose of the research so far as practical application is concerned. This served to motivate him to put his best effort into the practice. Then he was given a half hour or more of preliminary trials on cards not to be used later. The experimenter showed him where to look for the stimulus, and told him to leave from about a half second to a second of time after the stimulus to place his hands for the notes and then play them promptly without debating mentally whether he was right or wrong. If a finger slipped, the performer might repeat the passage to play what he intended. Under any other conditions, a hesitation would be marked as a mistake in time. The subject was directed to play in strict time at any convenient moderate tempo. Any introspections as to how the reading was done or why it was too difficult were recorded on the back of the data

sheet from time to time throughout practice. Any persistent errors were noticed and recorded as such on the sheet. The preliminary trials made the subject adapted to the experimental situation and the rules of the game before results were taken. A half hour usually fixed the simple habits needed for correct following of instructions. The subject found in this time the distance at which he could read best, which varied individually as it naturally should.

The actual experiment began. Results were treated more in detail than in diagnosis. All cards were given in all series except the last few in F, which were far too difficult for some, and two entire series, G and H, which turned out to be unsuited to tachistoscope work. Cards in other series that were too complicated to be read by most subjects in a short exposure were given anyhow in order to see if part of the material could be read.

Each subject practiced three times a week, an hour at a time, except during the summer session of 1937, in which the limited time made five practice hours a week necessary in order to finish the work before the vacation. No strain was reported by most subjects near the end of an hour, but three suffered occasionally from headaches after the practice hour was over. These individuals needed a change in their glasses, and suffered some eye strain in other activities. Unfortunately a scarcity of qualified subjects made the exclusion of those with serious eye defects impossible, but the writer believes that the effects of this factor were slight. Regularity in attendance was difficult to secure because of illness and the irresponsible attitude of many undergraduate students. Effects of absence from the experiment seemed to be only temporary, however, and a few minutes only were taken to readjust the person fully to the conditions. This irregularity did, however, cut down the number who could be obtained for 30 hours of work.

The rapidity with which the material could be covered varied individually, because some gave more introspections than others concerning how they read, and because some were prompt in response while others were slow and uncertain for hours in spite of their efforts to follow the instructions and react quickly.

Consequently the total amount of material read was not the same for everyone. About five readings of the entire collection of cards could be covered in 30 hours. Some of this time was taken for preliminary instructions and trials, introspections, specific instructions of what signature or clef sign to keep in mind for each series, the final test, etc. No one card was shown to the subject on two successive days. After seeing it once, he did not see the same one again for at least a week, and usually ten days. Consequently memory for any specific material would be expected to be slight. Not even the sequence of the cards was altered.

The subject was not told just what errors he was making, since wrongly read cards might remain in memory better than correct ones, but he was allowed to examine his record briefly after each series. When he saw general improvement on his data sheets, he was motivated to greater effort so that his progress might continue.

In order to check on the practice effects at the end to see whether they were specific or general, a T series was collected for testing. No subject had read the cards of the test during practice. All errors were recorded in order to compare the efficiency on these cards with that on the practice series. Finally the two Lalo selections were presented a second time without the tachistoscope. The subject had not seen these for over two months. Again a subjective evaluation of reading efficiency was entered on the data sheet by the experimenter in terms of the letter ratings described earlier. A comparison of the first and final readings of the two passages was then possible. The number and type of the errors made both times could in some cases be fairly accurately recorded after each reading. If the accuracy was too bad, this was impossible, of course, but some comments on outstanding deficiencies were made in every case.

The writer is fully aware of certain inadequacies in his subjective evaluation test, but he felt that the practicality of the tachistoscope technique needed demonstration to prove that the measured gain in span of perception and accuracy might be applied in the ordinary reading situation.

Following a suggestion given by Dr. W. B. Pillsbury, the

experimenter took a few photographs of eye movements made during the reading of certain types of music which could not be read very well tachistoscopically. Unfortunately no such elaborate equipment as that of Jacobsen was available for this purpose. Therefore the widely used Ophthalmograph was tried. All horizontal movements over a card bearing a simple melodic line could be clearly demonstrated, but vertical and diagonal ones, which would obviously be frequent in any elaborate music, were lost so far as this camera was concerned. The cards used for these photographs were illustrations of polyphonic music to be played on the piano by both hands, therefore the films did not show which melodic line was being fixated at any given moment. Projection of the pictures on the cards so that the image was equal in size to the card photographed gave the number of fixations made, and their location horizontally. The experimenter had no exact means of determining whether the reader was fixating a note in the right hand or one immediately below it in the left hand part, however. In some cases he could guess with fair certainty where the fixation was, but in others he could not. Therefore this approach was a mere side issue, and could not be carried out with success as extensively as desired.

IV. Results

1. *Individual Differences in Span of Perception.*—The fifty individuals who were tested for diagnostic purposes were divided into three groups. The first ten subjects were professional musicians, among whom are listed two outstanding, internationally known concert performers, who are now teachers in two leading universities, and other musicians of high rank. Thirty individuals have been included in the second group, who have had various amounts of musical training all the way from a conservatory diploma to a few months' study of an instrument only. Ten subjects who could not name the notes on the staff with certainty made up the third group in this experiment. Two groups are treated separately in Table 1, in which mean scores

TABLE 1

MEAN SPAN OF PERCEPTION AND RANGE

Professional

Subject	Training in Music	Melodic M	Melodic R	Polyphonic M	Polyphonic R	Chords M	Chords R	Type of Reader
1.	Organist	5.3	1–9	5.7	2–8	3.9	2–9	Pattern
2.	Pianist	5.2	2–9	5.8	3–8	3.6	2–9	Pattern
3.	Pianist	5.4	1–9	5.1	2–7	3.6	2–8	Pattern
4.	Violinist	5.1	1–7	5.1	2–7	3.6	2–8	Pattern
5.	Pianist	5.8	2–9	5.6	2–8	3.9	2–9	Pattern
6.	Pianist	3.2	2–7	3.1	1–6	3.1	2–7	Part, fills in subjectively
7.	Violinist	4.8	2–7	5.0	1–7	3.6	2–6	Pattern
8.	Violinist	4.0	2–6	3.9	2–6	3.5	2–5	Pattern
9.	Pianist	5.3	2–9	5.7	2–8	3.9	2–8	Pattern
10.	Violinist	3.1	2–6	4.0	1–6	3.1	2–6	Part, completes subjectively
Mean		4.72		4.9		3.59		

Non-professional

Subject	Training in Music	Melodic M	Melodic R	Polyphonic M	Polyphonic R	Chords M	Chords R	Type of Reader
1.	10 years	2.9	1–5	2.1	1–5	3.1	2–4	Part
2.	10 "	2.7	1–4	2.0	1–4	3.2	2–5	Part
3.	9 "	2.9	1–5	3.0	2–5	3.1	2–5	Pattern
4.	9 "	3.8	1–7	3.4	2–5	3.6	2–6	Pattern
5.	9 "	2.1	1–4	2.2	1–4	2.9	2–4	Part
6.	8 "	3.8	2–5	2.9	1–5	3.5	1–6	Pattern
7.	6 "	2.7	1–5	2.1	1–4	3.1	1–5	Part
8.	5 "	2.9	1–5	2.4	1–5	3.6	1–6	Pattern
9.	5 "	3.8	2–7	3.4	2–6	3.6	1–7	Pattern
10.	4 "	2.4	1–4	2.1	1–4	3.0	1–4	Part
11.	4 "	2.5	1–4	2.0	1–4	3.0	1–4	Part
12.	4 "	2.5	1–4	2.0	1–4	3.0	1–4	Part
13.	4 "	3.8	1–5	2.9	1–5	3.7	2–7	Pattern
14.	4 "	2.7	1–4	2.1	1–4	3.0	1–4	Part
15.	4 "	2.9	1–5	2.1	1–4	3.6	1–6	Pattern
16.	3 "	2.7	1–4	2.2	1–4	3.6	1–6	Part
17.	3 "	3.8	1–7	2.4	1–5	3.5	1–6	Pattern
18.	3 "	4.1	1–7	2.9	1–5	3.5	1–6	Pattern
19.	3 "	3.8	1–6	2.4	1–6	3.8	1–6	Pattern
20.	3 "	3.7	1–5	2.4	1–5	3.1	1–5	Pattern
21.	3 "	2.7	1–4	2.1	1–4	3.1	1–5	Part
22.	3 "	2.8	1–5	2.1	1–5	3.1	1–5	Part
23.	3 "	3.1	1–5	2.4	1–5	3.3	1–6	Pattern
24.	3 "	2.3	1–4	2.0	1–4	3.1	1–4	Part
25.	3 "	1.7	1–5	1.2	1–5	2.1	1–3	Part
26.	2 "	1.6	1–4	1.0	1–3	2.1	1–4	Part
27.	1 year	1.2	1–3	1.4	1–3	2.3	1–4	Part
28.	1 "	1.5	1–3	1.0	1–2	2.1	1–3	Part
29.	1 "	1.1	1–3	1.2	1–3	2.0	1–3	Part
30.	6 months	1.4	1–3	1.1	1–2	1.9	1–3	Part
Mean		2.73		2.15		3.06		

are given for each person for the three types of material used in the test. The scores represent the number of notes that could be read at one exposure. Naturally the range of these scores was wide, because some patterns were of greater complexity than others. Many factors operated in determining the number of notes that would be accurately perceived on each card. The area covered by the group of notes made a difference, naturally, since the notes that did not fall on or near the fovea were illegible. Hence, within limits, the closer the notes were crowded together, the more of them were seen at an exposure. Wide spacing has been used for quarter notes and narrower spacing for eighths or sixteenths in nearly all printed music for the simple reason that experience has proved that we can read them more easily this way. Faster notes must be seen in larger groups than the slower ones in a melody.

The mean scores for melodic material given in Table 1, therefore, were based on reading of close spaced groupings in every case, in order to make the conditions as uniform as possible for all individuals tested. The polyphonic material used in the test was two-part music for one hand, since music on two staves was impossible for the majority at that short exposure. The chords included in the test were in half of the examples on one staff, and in the other half on two staves. Since there was a wide variation in span of perception, some subjects had to be tested on three note groups, others on four, and still others on six. If all subjects had been given six note cards, confusion might have resulted when the less advanced performers attempted to get all that was there.

The range of the numbers of notes seen is given to show that some patterns were almost impossible without analysis, while others were so familiar that only a few cues were necessary to call out the complete response. Errors caused by looking too high or too low, etc., were not included in calculation of the span if discovered. Usually no notes were correct if this happened, and the observer was given another chance. Table 1 includes the subjects who could play at least the simple cards. Ten others were treated separately.

(a) *Totally untrained subjects:*

1. This individual could read only one note at a time with certainty as to its location. If given a melodic progression of two notes, he could tell in which space or on which line one note was located, but failed over 50 per cent of the time with the other note. In locating two notes on the staff on Series I, he scored 90 per cent success, but failed almost entirely to locate three note chords. The comparison technique applied to this individual gave about the same result as the description of location. The subject reported that he did not see the notes in groups.

2. The result was practically the same as for the preceding subject except that if two notes were on adjacent steps in melodic progression, she described their location correctly in the five cases of this kind. She reported that she picked out single notes, and could not describe the shape of patterns.

3. This individual reported that he could see configurations vaguely, but could not locate them on the staff, which he could not see as clearly as the notes. Comparison cards showed him to be very uncertain of the differences in details of even the simplest patterns.

4. This subject scored only 60 per cent success in locating one note on the staff. He reported that he did not see groups of notes as wholes, but attempted to pick out one note.

5. Though he could not play or name notes rapidly at all, this man has sung by ear, being guided by note position. He made a mean score of 1.5 notes for melodies and 2.0 for chords, telling promptly the exact location of the notes he saw clearly. In larger configurations he observed direction of motion only, and could not give details.

6. The response was inaccurate even if only one note was shown, but the subject occasionally located three notes correctly on the staff. He claimed to see the group as a whole in the cases in which he succeeded, but he failed in over 50 per cent of the trials.

7. The report was similar to subject 6, except that in this case the accuracy of location for chords was greater than that for melodies.

8. This subject saw groups of three notes as wholes, according to his report, but had difficulty in seeing the lines of the staff. Therefore his descriptions of the location of the notes were frequently wrong because he placed the entire pattern too high or too low by one line or one space, even though he grasped the relative position of the notes accurately. He located one note correctly 90 per cent of the time, but his failures greatly increased if two notes were shown.

9. This subject scored higher on chords than on melodies. His mean on the former was two notes, and on the latter 1.5 notes. He could not give details of the patterns of three notes or more.

10. The report was identical with subject 9.

An examination of Table 1 reveals the fact that in general the greater the amount of training, the better is the reading of these cards. There are a few noteworthy exceptions that will be discussed later. Polyphonic music naturally confuses those who have not had sufficient training to have mastered its difficult points through experience in playing it. The years of training in the second group are the number of years of piano study. Public school music was not counted, nor was voice training, of which four subjects had a little. The last column contains information as to whether the subject usually grasps a group of three or more notes as a pattern, or whether he sees only one or two single notes which serve as cues, and completes the group subjectively. This fact was brought to light not by introspection alone, but chiefly by the nature of the response and the errors made. Those who reported that their impression was of the whole responded, with but few exceptions, without hesitation a brief fraction of a second after the exposure. If they made mistakes, these were generally errors of locating the entire group too high or too low on the staff. Those who reported that they saw only one or two notes clearly demonstrated the fact that they had no idea of the shape of the pattern in that they usually got one note right, but merely guessed at the rest. Working out the part that was illegible to them by logic took at least a second after the

exposure, often much more time. When the response finally occurred, there was a noticeable uncertainty about it in contrast to the prompt, self-confident response of the pattern readers.

Naturally the line of distinction between part and pattern readers is difficult to draw in some instances. For the very best subjects some patterns were impossible to grasp. Only parts of these were seen, and completion had to be subjective. The slowest readers rarely see whole groups, but exceptions to their part method do occur if the groupings are easy. Therefore the pattern response to notes seems to be a matter of degree, and our distinction here is arbitrary perhaps, but it may prove practical. At least it is closely analogous to Dearborn's A and B classifications referred to in Chapter II. Whether anything corresponding to Dearborn's C group occurred in this experiment is as yet unknown to the writer. Further research may clarify this point.

Subjects who have studied harmony and counterpoint extensively scored higher, as a whole, than those who had studied only an instrument, but here again there were noteworthy exceptions. Men did not do significantly better than women, though not enough subjects were used to determine whether a small, but reliable difference does exist. Experience in playing the music of Bach invariably raised the efficiency in reading the polyphonic selections from this composer, even though none of the subjects were familiar with the selections to be read. Experience with ultra-modern music aided, as might be expected, in the recognition of certain chords used in this experiment which others not having such experience failed to get.

(b) A consideration of the individual results of *professionals in the first group* might be of interest. Subject 1, a concert organist and teacher of high rank, is noted for his facility in reading the most intricate passages for organ or piano almost without error the first time. The rapidity with which he takes in three or four melodic lines in reading a fugue or chords in even a quite unconventional progression has amazed many of his colleagues. He averaged above five notes on melodic and polyphonic material in span of perception, but was not as high on chords because most of the examples of this type contained only

four notes or less. His response was made less than half a second after the exposure in all but two cases. He found card B 1 puzzling (see Fig. 2), and saw only the first note. He read six note groups in series D without difficulty in estimating the wide intervals, unlike less experienced readers. He responded to 9 notes of D 43, but did not see the last 3 notes. Examination of this card (Fig. 2) will reveal that mere surmising was highly improbable in this case. He reported that he got only a few essential cues in reading K 25, in which he read all of the nine notes correctly. His introspection follows: "That time I did not see all the notes clearly. I saw the general form of it—that it was a G major chord, you see, and then, just by their distribution, I guessed what the rest were most likely to be, and happened to get them right."

Because of his interest in modern idioms, this organist read the chords in the latter part of L correctly, using their general form as a cue to recognition. Few can achieve this. He reported that in Series F he encountered several forms that did not register, while others were comparatively easy to grasp. His errors as well as his introspections revealed the fact that he nearly always used the general form as a cue to aid him in recognition. He pictured E 1 as a scale, for example, D 4 as a broken G minor chord, F 25 as an ascending passage of thirds, L 3 as an augmented triad, etc. He did not seem to start at the bottom and read up, or even to do the reverse, but in most cases saw the whole at a glance even if it was on two staves.

Subject 2 in the professional group was also noted for efficient reading. She is an internationally known concert pianist. Her reading seemed to be definitely by general form, though she was not clearly conscious of how she achieved it. Card B 1 also gave her difficulty, while the six note groupings came easily, with a few exceptions. The first two subjects found exactly the same cards to be the most difficult.

The third performer, a pianist, also used the form as a cue to recognition, but was not quite as prompt in responding as the first two. Though technically efficient and gifted in music, he is known to be less accurate than the first two subjects in sight

reading. The difference is not great, however. He tended to depart from his usual procedure in response to modern idioms. He could not integrate unusual chord patterns well, and tended to work them out by logic from a few cues, making them conform to some conventional chord pattern closely resembling the chord presented to him. His reaction time exceeded two seconds on many of these, indicating a possibility that he deciphered such combinations from a clear photographic memory of them. He was rarely conscious of such a procedure. He reported that the presence of fingering marked on the music frequently helped him to get the notes. All other subjects appeared to ignore the presence of these markings when reading, and some reported that they never saw any. Only one was distracted by the presence of fingerings.

Number 4 is a concert violinist internationally known, and also a teacher of violin. He plays the piano, but not as efficiently as the three preceding people. One might expect him to be higher in mean score on melodies than pianists, and comparatively less efficient on chords, but this is not the case, as can be observed from Table 1. He is an efficient reader, even of the most difficult material for his own instrument. Some of his introspection on what his difficulties were read as follows: "I can see the notes all right if they are solid, like quarter notes, but if they are open notes, they are not as plain. Also I notice I have trouble seeing the staff. The lines are not plain to me like the notes are. Sometimes I don't see any lines." He frequently expressed surprise at an unexpected change in position of the notes from one card to the next. He complained that many of his errors were the result of having no chance to anticipate what was coming next. This complaint was quite common, and it brings out one essential difference between the experimental situation and uninterrupted reading. He appears to use general form as a cue, but is unconscious of the details of his procedure.

Subject 6 is an efficient teacher of piano and a fair performer, but a slow, stumbling reader in spite of extensive training. Her responses show that she did not notice the general form of a group of notes, but saw only two notes clearly. She was clever

at finishing patterns subjectively, and frequently succeeded, but occasionally revealed the fact that the subjective was playing a larger part than it should. Her reaction time was frequently two seconds or more, and her response was likely to be hesitating unless the material was easy. Wide skips bothered her in A 5, A 37, C 4, and C 7 (see Fig. 2). She tended to underestimate these intervals.

Nothing particularly striking was found in the records of the next two subjects, except that they had the greatest difficult with the same cards that troubled all of the others.

Three musicians of the professional group, 6, 8, and 10, could not read both hands in Series J, K, or L with greater than 40 per cent success. They could get one hand or the other in a chord, but looking midway between often resulted in missing both parts. The others saw all the notes clearly on both staves if they were not spread out over too much space. Universal difficulty was experienced in perceiving two-part polyphonic music for both hands at a short exposure. Probably jumping back and forth is necessary, and the maximum number of notes per fixation of the eyes is three in this kind of music as a rule. Few exceptions occurred, and doubtless the tachistoscope is inadequate to measure performance on polyphonic material for both hands. However, the experimenter found his results interesting and worth while on polyphonic music.

(c) A survey of the results in the *non-professional group* reveals the fact that at all stages in musical education one can find inefficient readers and skillful ones. Talent appears to have little to do with it, since teachers often report that some of their most promising pupils are unable to make much progress in reading, while others read well, but have different limitations. A few examples will illustrate special disabilities.

Subject 1 is a fair performer on the piano, but reads slowly. Her accuracy was fair on the test, but her span of perception was limited. She was inclined to ponder over a stimulus card five seconds while looking at the floor. Then she played two, or rarely three notes with self-confidence, as if she was certain of the response. She was sure that she was right, and usually

she was correct as far as she went, but if the pattern included more than two or three notes, she was unable to obtain any accurate impression even of the direction of motion of the rest of the notes. There was no evidence of any anticipation in her reading, and her thought processes were slow in reading music in spite of thorough theoretical and practical training. She was greatly confused by polyphonic music, even on the same staff, and could not read chords for both hands at a single exposure. This case is an example of a limited span of perception together with slow habits of thinking, an over-careful attitude, and too many inhibitions.

Number 4 is an example of a good reader, who plays well. He reads by patterns and responds quickly as a rule, but he has some individual limitations. Wide intervals bothered him. Any skip of more than a fifth in a melodic pattern caused inhibition of his usual prompt response. After two seconds or more he either played something to represent the shape of the general configuration, missing some notes by one step usually, or else remarked, "I've lost it now." He was inclined to overestimate wide skips by one or two steps. He grasped large units of material of a conventional sort, but failed miserably with ultra-modern chords, because of his lack of experience with that type of music. He tended to make these conform to conventional chord patterns, frequently admitting that his response was a mere guess. He said that these peculiar chords did not make sense to him as the conventional ones did. His record clearly showed that major, minor, augmented, and diminished triads in any inversion or position were meaningful units to him as were also most seventh chords unless altered. Ninth, eleventh, and thirteenth chords were usually spelled out one note at a time mentally. When he finally played what he had figured out, it was usually less than half right.

Not much difference was found between the results of the typical professional musician in the first group and the record of subject 13 in the second group, who reads exceptionally well, even though he has studied only four years. He plays mechanically, without much interpretation, and does not consider himself

musical. He has little theoretical knowledge of music, and his piano teacher did not consider him talented. He can read the correct notes in time of almost any composition within the limits of his technical ability, even in the realm of the ultra-modern to some extent, though he claims that such music does not make sense to him. His correct readings of unusual chords were more frequent than one would expect, considering his limited training, because most of the subjects of his degree of advancement did not read this type of material in patterns. They picked out a few cues and almost always failed. The writer does not know of a satisfactory explanation for this exceptional case. As a general rule, experience with a given type of material increases the span of perception for it, but this subject had not played polyphonic or modern works one-tenth as much as many more advanced performers, yet he excelled them in reading both types. He did not remember anything in his musical training which would have caused him to form the habit of grasping tone groups. As far as he knew this habit came about by mere accident.

Table 1 does not show some of the changes that were evident in the performance as those of lower and lower degrees of advancement were tested. Beginning with the students of four years and increasing as the amount of training decreased was a tendency to count leger lines above or below the staff instead of responding immediately to music so located. Professionals and most of the advanced students are usually able to read notes not more than three leger lines above or below the staff without counting these lines, but the less advanced failed on numerous cards because they lost the impression while trying to count. Occasionally clear imagery aided in retention of the visual pattern until a response could be counted or reasoned out. Subjects 16 and 19 showed evidence of photographic memories, which made up for the lack of automatization of their responses to the printed notes by playing them upon the piano. Number 19 was skillful with his fingers, but had played so much by ear and so little from the page that he was slow in deciding what a note was, even though he retained a clear impression of its location

on the staff. Subject 16 reported that she was always puzzled when she saw accidentals, and had to stop to think them out.

Among those of two or three years' training, there were several whose excessive playing by ear resulted in a lack of automatization of response to individual notes. Number 24 had difficulty in finding notes on the bass staff, because she was more familiar with the treble staff on which all the notes she sang were written. She was in the habit of guessing the bass, not really reading it, because she could figure it out by ear in simple music by finding what would sound best with the treble. Even the beginner does much surmising in reading music, as he does in reading words, but he is not as successful with his subjective completions as is the more experienced performer.

Examination of the table of results for the professional group will reveal that the average span of perception is greatest for polyphonic music, slightly less for melodic examples, and considerably less for chords. There are small individual differences in the proportions of these figures, but the general rule holds good that fewer notes were correctly perceived in the harmonic material. The cause of this lies partly in the difficulty of reading on two staves at once, one usually being bass and the other treble. Most of the polyphonic and all of the melodic examples were on one staff. However, it is interesting to note that in the second group the mean scores of all 30 subjects are reversed. Chords now rank highest, while melodies come second and polyphonic cards third. It is easy to understand why the polyphonic type is extremely difficult for those with less training. Success with it falls down more than that for other types as one reads down the table. However, it is more difficult to discover why the chords are easier to read than the melodies for many of the less efficient readers. The explanation may be partly in the nature of the particular examples used in the test. Some of the notes in the chords were crowded together more closely than any of the notes in the melodies. Most of the three-note chords were on one staff. The four-note examples often had three notes on one staff, or even all four in a few cases. Therefore for pur-

poses of comparison, the number of correct readings of D F A written as a chord on the treble staff was checked against the number of successes with the same notes given in melodic succession on the treble staff. Out of 40 readings of each, there were 29 successes with the melody and 36 with the chord. All four of the subjects who failed to read the latter correctly missed the former also. The most common error with the chord was to play it one space too high, that is to say F A C instead of D F A. The most common error with the melody was to mistake it for a stepwise progression, D E F, a sign of extremely inaccurate observation of details. There were three subjects who played D F G, which indicates that the last note probably lay outside the field of clearest vision, or else was forgotten while the first two were being figured out.

Other examples were compared to check on this point further. The fact that successions of two or three notes were more difficult than the same notes written under each other in the form of a chord was illustrated in 8 examples. Thirty subjects who read these examples averaged 7.8 right for the chords and 6.2 for the corresponding melodic patterns. Individual differences were not very great, and the same tendencies to error were universal. In the case of the chords, the subjects tended to get the relative position of the notes with respect to each other quite accurately, but failed to place the pattern correctly on the staff in a few cases. They reported in their introspections that they saw one-hand chords plainly, every note being as clear as every other note usually. Exceptions to this were rather rare. However, they reported that their difficulty was chiefly in determining which lines or spaces the chord occupied. One of them, who was well acquainted with psychology, expressed his view as follows: "The notes stand out, but the staff is not a part of the figure. It is background to me, and not noticed so much. The notes attract my attention, but the staff does not. It seems to me that there is a figure-ground distinction there in which the staff is ground." Following this up, the experimenter questioned two others acquainted with the Gestalt concept of figure and ground, and for them the staff was a part of the figure.

The outstanding tendency in the corresponding melodic phrases was to miss the last note by placing it too high or too low by one step, or more rarely by two steps. The probable cause of this error was failure to see this last note clearly, since the introspection indicated that the first note or two notes appeared to be much plainer than the last note in many cases.

Another factor that is of importance in the span of perception is the area covered by the pattern. The horizontal distance along the staff over which notes were clear enough to be legible was roughly determined from the results. Since ordinarily most errors occurred at the end of a group, the experimenter could measure the length of the part accurately read on cards on which correct guessing was unlikely. From this he could determine with fair accuracy the width of the legible portion of the field. This distance, which was calculated for every subject from five cards, ranged from about 3/8 inch to 1 3/4 inches, and the mean was approximately 7/8 inches. The vertical distance over which chords were legible was similarly estimated. The figures ranged from about 3/8 inch to 1 1/4 inches, but the mean was again approximately 7/8 inch. Unfortunately the distance of the card from the eyes was not recorded, but observation indicated that those who set closer to the apparatus than the average covered a smaller area than the ones who sat farther back, as might be expected. Nevertheless under the conditions of this experiment, any change from the accustomed reading distance of each subject was not advisable, especially for those who practiced reading with the tachistoscope for many hours, because such a change would result in eye strain and decreased clearness for many people.

2. *Analysis of Individual Difficulties.*—We turn now to a consideration of the more thorough study made of 24 individuals not included in the groups discussed in the previous section. As stated in the description of the practice experiment in Chapter III, these people were given all of the material to read except that which proved entirely too difficult for them. They constitute the practice group.

The knowledge that could be obtained about their individual

disabilities was much more extensive than that from the diagnostic group, although none of the results conflicted on any important point. There were several difficulties that were common to a majority of the subjects as they went through the various series, and these will be discussed first.

Twelve subjects claimed that although the manuscript cards were plainly written, small irregularities made them more difficult to read than printed notes, although the difference in difficulty was small. The remaining number did not notice any difference. The record for the entire group for Series A, which was in manuscript, indicated little loss of accuracy for this reason. The wide skips of a sixth or more were more frequently missed than the smaller intervals, *i.e.,* fourths, thirds, and seconds. Octaves were more easily perceived than sixths, sevenths, and ninths. The latter were frequently mistaken for the former by considerably over half of the subjects. The introspections that were recorded on the difficulty with wide intervals indicated that usually one or the other of the notes was not seen clearly, which suggests that possibly in uninterrupted reading these combinations are read with two fixations of the eyes, one for each note.

Series B, in the first half, gave all subjects persistent trouble, because at that short exposure they could not grasp the time relations on some of the cards. Eighth notes, unless connected together by a bar, were unobserved on a number of cards and played as quarters by 14 subjects. The remaining 10 seldom had this difficulty but they occasionally failed to observe which notes were eighths and which were quarters. The frequent failures on the last note were caused by the fact that the distance covered by many of the patterns was too great to be seen clearly at a glance. The clearness with which the notes in the last part of Series B were seen was more nearly equal for the three notes, because they were confined to a small area. Therefore few failures on the last note occurred. Nearly all subjects made occasional misplacements of entire patterns on the staff. The less efficient readers in the group, who numbered eleven, frequently played B 21 as B C ♯ D instead of B D F ♯, showing that they observed only the first note and the direction of motion

of this group. Seven of these played B 22 as G F ♯ E, showing a similar inaccuracy in observing details. These and other errors of the same kind persisted for three subjects through two or more trials, but for the others accuracy increased with practice, eliminating this difficulty by the second trial. There were two or more instances of failure to remember the signature of the last part of Series B on the records of thirteen observers.

The records on Series C showed that complexity varied not only with the number of notes in a given pattern, but with various optical and musical factors. Some cards were difficult for everybody, while others gave trouble for comparatively few. For example C 3 was persistently missed by only 2 subjects, while C 4, C 7, and C 19 caused some failures on every record (see Fig. 2). The distinction between the easy and difficult patterns, from these and other examples, was found to depend upon the width of the intervals between tones, the consonant or dissonant effect of the intervals, and position on, above, or below the staff. The wider the skips, other things equal, the greater were the chances of failure. Observers tended to make note groups conform to a chord or scale pattern, and end in such a way as to produce a pleasant effect. As mentioned before, the use of leger lines, as in C 19, resulted in counting, and consequent loss of the impression.

For similar reasons there was a gradation of difficulties in Series D. With these longer groups, a possible subdivision, on the basis of arrangement of the elements, into two groups of three, or more rarely three groups of two, greatly facilitated the response. Such was the case in D 1, 2, and 3, but not 12, for example. D 4 was easy because it followed a familiar broken chord pattern, but D 12 was impossible for most subjects to grasp at all.

Frequent modulations in Series D caused a greater number of failures to remember the one flat in the signature than might be expected. Fifteen observers failed persistently in this respect until near the end of their practice. The four sharps in the signature from D 41 on caused some failures for 16 subjects. As a general rule the less highly trained individuals had the

greatest difficulty keeping signatures in mind. The added difficulty of the presence of more notes than could possibly be seen clearly at a glance caused forgetting of sharps in the signature as well as confusion with regard to time values. The part readers were invariably confused, and played one or two notes or failed to respond at all. They had no clear idea of the direction of the progression, even if they had played the first two notes correctly. The pattern readers in the group grasped 3, 4, or more notes, usually the first part of what was on the card. The four best readers were able in addition to get the direction of motion of the rest of the material on the card. They often failed to observe the exact notes on the part vaguely seen, but they could indicate by their responses that they were looking ahead and anticipating what was to come next in order to prepare for it. For example these best readers gave an accurate rendition of the first five notes of D 41, but only vaguely anticipated the last six notes. They knew that these ran up, but how far they did not know. Two of them saw that these were sixteenth notes going up the scale, but nothing more definite about them than that. The irregularity in the middle of the progression on card D 43 was either unobserved or, in a few cases, seen too vaguely to be attempted.

The arrangement of E 1, 2, 6, and 7 (see Fig. 2) is such that a greater number of notes may be included in the span of perception than in Series D. Obviously accurate perception of details is not so necessary where irregularities are absent. Ten subjects persisted in playing one or two notes too many on each card. Eleven frequently misplaced the entire progression one space or one line too high or too low. E 9 was persistently played wrong by 20 subjects. The most usual mistake was to play five notes in each group: G A B C D, failing to observe that one step was skipped. Another common error was to skip the wrong step, for example G A C D. The E flat on card 24 was unobserved by 12 subjects, who persistently played the two groups alike. E 25 was similarly mistaken for two identical groups, but E 26 was played, with but one exception, with the second group unlike the first, but seldom correct.

Polyphonic music for one hand was more difficult than a single melodic line, but two independent melodies in conventional counterpoint were read simultaneously by efficient readers. Errors in the lower part were nearly twice as numerous as those in the upper part. Introspections showed that most observers regarded the lower melodic line as subordinate, and of lesser importance than the top line. Parallel motion was more easily grasped than contrary motion of the parts, and like time values were nearly always played in time, while two notes of one part against one of the other caused considerable confusion. Memory for the four sharps failed persistently in nine subjects. Card F 27 was missed entirely three times as often, approximately, as F 25, because of the fact that consecutive fourths were comparatively rare in the playing experience of nearly all subjects.

The result of attempts to read two-part material simultaneously for both hands was a large percentage of failures. Jacobsen's results indicated that this kind of material is not read by a series of units taking in both hands at once, but that shifts of the eye must be made from one staff to the other. The best subjects in this research, however, were able to read simple cards of this sort fairly accurately in the tachistoscope at a glance. Whether or not their usual continuous reading involved such a procedure, taking in patterns including both hands at once, was unknown to these individuals. Therefore a series of eye-movement photographs taken as described in Chapter III helped to clear up this point.

The results of the projection of these films upon the cards read gave the number and approximate location of the successive fixations of the eyes made in reading the card. Figure 3 shows diagrammatically the manner in which a few patterns were read. The location vertically of the pauses was uncertain in a few instances. The writer may not have guessed right in these cases as to whether the subject was fixating the left hand part first and then the right hand just above it or the reverse. Occasionally the horizontal location of the fixations and of groups of notes offered helpful clues, but the writer does not consider these results to be sufficiently accurate to warrant further study of eye-move-

ment records without more adequate apparatus for the purpose. These few diagrams demonstrate conclusively, however, that the readers who were tested frequently shifted their gaze from one staff to the other, taking in a note or a little group of notes at each fixation. They were able to take in two-hand combinations at a glance when these were simple, but did not follow such a procedure in continuous reading of the same material.

A few more photographs were taken besides the ones represented in Figure 3. They demonstrated a similarity in the procedure of E and that of W, both of whom are good readers, in reading the same material. However, results on a third subject, who was a slow and stumbling reader, were too complicated to make accurate locations of his fixations possible without better apparatus. The behavior of his eyes gave evidence of unsystematic procedure, however. There were many regressions on the film, and twice as many fixations as were made by either E or W.

The results on chord reading indicated that the most efficient readers do not spell out a chord, but read it as a whole, unless it is too complicated. A chord that is spread out often covers too great a vertical distance to be perceived without shifting the eyes, and in such a case only the part for one hand is seen as a pattern. Which part will be seen in the tachistoscope in such cases depends upon the direction in which the subject happens to be looking. Under usual conditions, however, the good reader was able to read two-hand chords accurately at a glance. Whether or not this is the procedure in continuous reading remains to be seen. The subjects themselves are uncertain about whether they read chords up or down. Five of them reported that they did not read chords up or down any more than they read single words to the right or to the left, but that they grasped the entire unit almost instantly in most cases. Their accuracy on the harmonic material in this research indicated that their statements were probably correct.

Ten of the less efficient readers reported that sometimes they read chords up and sometimes down. Four of the most advanced players among the part readers gave clear evidence of spelling

chords up, while two beginners gave equally good evidence of always spelling them down, beginning with the highest note in the right hand. The remainder were irregular, and used no consistent procedure.

Checking the comparison series against the other technique for each kind of material proved that if an observer could give the exact differences between details of similar patterns, he was accurate enough in his perception to reproduce the music on the piano, and in only two cases was this reproduction response so difficult as to cause the performer to make errors frequently and become confused by factors other than perception of what was on the card. In other words the level of difficulty of material that could be done accurately was, in all cases but these two, the same for the one mode of response as it was for the other. The response of playing the cards was reported by nearly all as being always easy. If they saw what was there, they could play it as easily as they could pronounce a word or a phrase. Playing the music gave the experimenter a more thorough record of the perception of all details of every pattern than comparison reports of *alike* or *different* and one detail of difference in the case of unlike stimuli. Description of what was seen on the cards was given after the judgment of *alike* or *different,* following the suggestion of Wertheimer, but this description proved more difficult than piano playing for most subjects, because during the slow process of telling the experimenter every note, the last part of the pattern was lost before the subject had finished describing the first part, or one stimulus was forgotten while describing the other.

An example of a typical comparison response will make this clear. The exact words of subject N were taken down as she responded to a card in Series N as follows: "They were different. The first was E, then E an octave higher, and C and A, and the second was—Oh! I had it a minute ago, but I can't remember. Anyway there was a difference near the first." She demonstrated greater efficiency at the piano, since she could respond before the memory impression was lost.

Another difficulty that was common in the comparisons was the fact that frequently the subject would fixate higher in one

window than he did in the other, thus receiving the impression of a difference in position on the staff when no such difference existed.

Material for one hand was easily read, but two-hand combinations were more difficult to describe than to play. Differences were not accurately located by any but the five best readers, but some of the less efficient readers could get vague impressions that these patterns were alike or different without getting very many details clearly. With this type of material as in the preceding there were patterns which seemed impossible to integrate, and others which were easy.

There are a number of individual records which deserve analysis in order to demonstrate the effects of training under different teachers upon reading. Subject A (see Table 2) is a violinist who has considerable technical ability, and has studied extensively both theoretical and practical music, but she spelled out her music one or two notes at a time. Her accuracy was fair at this slow rate, however. She did not stop to think before responding as a rule, but read only a few notes of a pattern, usually two, and guessed the remainder, which she admitted she did not see at all. She was most likely to see the first part of a pattern and have no idea of the shape of the whole, and frequently she got only one or two notes in the middle, and made errors on the first and last parts. She was persistent in this analytical approach, but finally became a pattern reader. She was not troubled by key signatures or by time values. For a person as well acquainted with harmony and counterpoint as she was, the number of correct anticipations of material not clearly seen was surprisingly small. This was true in several cases, suggesting that even with the most conventional type of material. guessing is not likely to be successful. She had difficulty with the lower melodic line in the two-part polyphonic material. She seemed unable to perceive immediately how the two melodies fit together. Therefore she played the upper one, and filled in a few notes of the lower. If the two parts were not on the same staff, she saw one or the other. The ultra-modern chords were not integrated well as wholes even after much practice, but were usually made

to conform to some chord that sounded more consonant in effect. The delay in the reaction was more than two seconds in these cases.

Subject G had the peculiar habit of perceiving the notes first and then looking back to get the accidentals afterwards. Sharps or flats in the signature were carried along as a part of the pattern, and were seldom forgotten, but added sharps, flats or naturals were often overlooked completely. When asked concerning them, she replied that she did not see any or that she saw a sharp or a flat, or something somewhere, but was not sure what it was or where it was located. In such cases her observation of the notes was usually accurate, even if there were six of them in a difficult arrangement. She reported that she did not know why she did not see the accidentals as a part of the pattern, but she knew that she added them on afterwards. She had no idea of the origin of this bad habit. Her reaction to simple cards, which was usually immediate, was delayed two seconds by the presence of one sharp or flat. When the response finally occurred, it was usually correct, unless the accidental was placed on the wrong note. Two accidentals frequently caused complete confusion. The reading of ultra-modern chords was much more difficult for this subject than the reading of conventional chords because of her failure to grasp accidentals with the pattern, though her perception of the notes was usually correct for both hands. Nothing in her past training, as far as the experimenter can find out, could account for her difficulty, since her experience with all types of music has been more extensive than that of others who experienced no such trouble.

Subject H has studied extensively under well known teachers. He is considered gifted, performs well, and has a phenomenal memory. He has memorized everything with great ease since childhood, covering a broad variety of repertoire including many contemporary works. He spelled out his music with accuracy, but slowly and never in time before he began tachistoscope practice. His teachers attempted to rid him of his disability by making him play a large amount of easy new music daily. This had practically no effect. His reaction time with the material of

this experiment was typically slightly over two seconds. Constant repetition of directions to play up to the given tempo was necessary. Hesitations were frequent. Sharps, flats, and naturals were confused and misplaced. He could not read a chord for both hands at a glance with certainty, but he did much intelligent guessing, some of which served to improve his record. The number of cues he actually perceived was too small. His recognition of time values was very poor. He showed no evidence of anticipating what was ahead. A quotation from his introspections, which might be of interest, is as follows. "If I once hear how it sounds, I can play it on the piano easily, but when I see those notes up there, I have to stop to think how they would sound before I know what to do with them. That is because I played by ear too much I suppose, especially when I was younger. Now I am trying to play by note, and I see how much I depended upon my memory before."

Subject R is a child prodigy of fourteen, who plays an extensive repertoire of medium difficulty or better from memory with finished technique and artistic interpretation. She is considered gifted by recognized teachers. Her record is not consistently good, because of unavoidable distractability and lapses of attention at times. When motivated to her best, this child read some patterns of fair complexity, but the number of patterns that failed to register as wholes was greater than that for more mature subjects. Her procedure in reading might be described as unsystematic, and her results were inconsistent. She failed most on modern idioms, with which she was totally unfamiliar.

Subject W is not musically inclined. His teacher did not consider him gifted, and he had no great liking for music, but he studied the mechanics of piano playing until he could play material of fourth-grade difficulty skillfully, but with little expression. His reading is phenomenal for a person of so little musical training. He read polyphonic material with amazing accuracy, considering his extremely limited experience with it, and ultra-modern chords which had to be spelled out by many more advanced performers registered in his experience immediately as wholes. The reason for his efficient reading could not

be discovered by questioning him about his training. He had none of the inhibitions in reading that were characteristic of many other subjects. He was dependent upon the printed page whenever he played, and could do little without it. This constant performing from the page was characteristic of the best readers among the subjects, with a few exceptions, and the majority of the less efficient readers memorized with ease. However, the writer finds insufficient evidence that an inadequate musical memory insures good reading, or that the performer who memorizes quickly necessarily reads badly.

Subject X shows clear indications of an emotional complex centered around practicing from the printed music. She was forced when a child to study piano with a severe and exacting teacher, who scolded her endlessly for every wrong note that she played, pointing out the correct note on the page. The pupil hated this teacher, and also printed music, which in her mind was associated with the teacher, but she enjoyed playing by ear and listening to good performers other than her teacher. She reported that the mere sight of a sharp, the cause of her most frequent errors at lessons, aroused extremely unpleasant emotions, even years after the lessons had ceased. Confronted with the task of overcoming many inhibitions in her reading in order that she might continue her study, she was motivated to work hard on this experiment, but the habits were persistent. Prompt response was blocked by a constant fear of making errors. Indecision on the simplest material resulted in loss of the impression because of delay. Much work was done before any patterns could be perceived clearly enough to play details accurately. Response to single notes was not well automatized, but this factor was a minor one in her case.

Thus it is evident that the individual background of every observer affected his results in an interesting way. These selected examples demonstrate the variety of music-reading disabilities, and the complexity of such a process when subjected to detailed analysis.

3. *Improvement With Practice.*—The mean span of perception is based on more data in this experiment than in the diagnostic

group, but the manner of calculating it was essentially the same. It represents the maximum number of notes that can usually be read by the subject with each kind of material. The range merely indicates the variation in difficulty of the music. Some patterns are impossible to integrate for most subjects, while other relatively long ones are easily read as wholes. The experimenter realizes the incompleteness of his tables, but in attempting to make the results concise and meaningful, he found elimination of minor details necessary. The most important of these details will be treated in the discussion later. The subjects are listed in order of their musical training on their major instrument (not always piano). Degrees are also indicated.

The subjective evaluation of continuous reading of each subject, which the experimenter made before and after practice, is recorded in terms of letters, as stated in Chapter III. The letters correspond to the typical college grading system, but the standards of performance for each letter were the same for everyone regardless of the amount of training or musical experience. Some attempt was also made to evaluate the piano-playing ability of each subject.

Only trials that were complete enough for accurate measurement were included in the table. If the last trial was only half completed, no computations from it were made. Limitations of the time that some subjects would work made equalization of the number of trials impossible.

Under *time* the experimenter placed the per cent of the total number of cards played correctly as to time values. Since only a limited number of cards were difficult in this respect, the percentages were fairly high. Hesitations were counted as time errors.

Under each type of music was given first the mean span of perception and then the range, as in Table 1. The material selected to obtain these figures consisted of 100 cards for melodic, 40 for polyphonic, and 100 for harmonic material for each subject. The same material was used for calculations in all trials for any one subject, but the same material was not used for all subjects. Each reader was tested at the level of difficulty at which he did best. Material that was far too easy or entirely

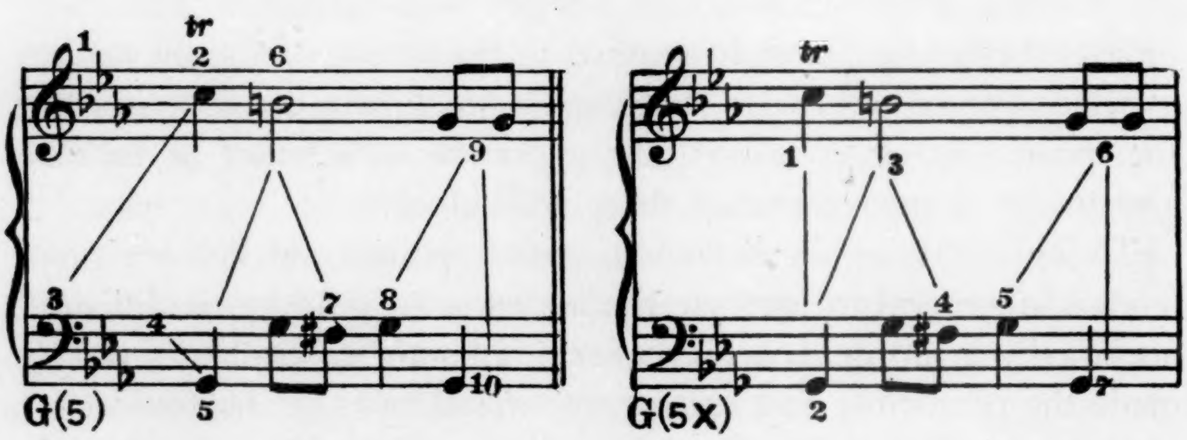
too difficult for a reader yielded inferior results; obviously, therefore, though it had been practiced, it was not included for purposes of measurement of progress.

In the last column is given the manner in which the majority of the cards were read during each trial. Every subject read some cards by recognition of the form of the entire pattern, and other cards by picking out one or two notes as cues, filling in the rest subjectively, but the dominance of one or the other method was obvious in nearly every individual. Some changed their manner of recognition on a large portion of the material as practice proceeded. This change was clearly demonstrated by the nature of their errors, and was recognized and reported in several cases by the subjects themselves.

A question may arise as to how the evaluation of a subject's reading at the end of a long period of practice may be compared with any degree of accuracy with one made at the beginning. To aid on this difficult point, the experimenter wrote down as nearly as he could the location and type of errors made by the subject as he played, keeping this record for comparison later. With extremely poor readers, this record could not be exact, but when errors were few, it was reasonably accurate. In every case of material improvement, the subject was aware of his own increased efficiency in sight reading at the piano. In eight cases, in which the subjects were studying music during the experiment, the music teacher confirmed the experimenter's belief as to the amount of improvement of these individuals.

Though training in harmony, counterpoint, and composition tended in general to increase the accuracy of reading, it did not necessarily insure high achievement. In some cases theory work made the performer so exact as to every detail that he could not attain any speed at all. Doubtless the acquaintance with rules of harmony aided in the anticipation of resolutions just as knowledge of grammar is helpful in rapid comprehension in book reading. Therefore more intelligent guessing might have been done by those with the theoretical background in the conventional sort of music. There were many without such a background, however, who demonstrated considerable ability to anticipate some

Numbers indicate approximate location of fixations made by subject E as he read the music. Arrows indicate the direction of motion of the eyes. In G 3 E made an error on next to the last note for the left hand, reading it as D. Accuracy was good otherwise.



Similarity in the reading of G 5 by subject E (left) and subject W (right). The signature was given before the exposure of 2 sec. Notice the irregularity of procedure near the first of E's reading where he tried to orient himself as to clefs and signature. Both readings were correct. Numbers indicate successive pauses.

FIG. 3

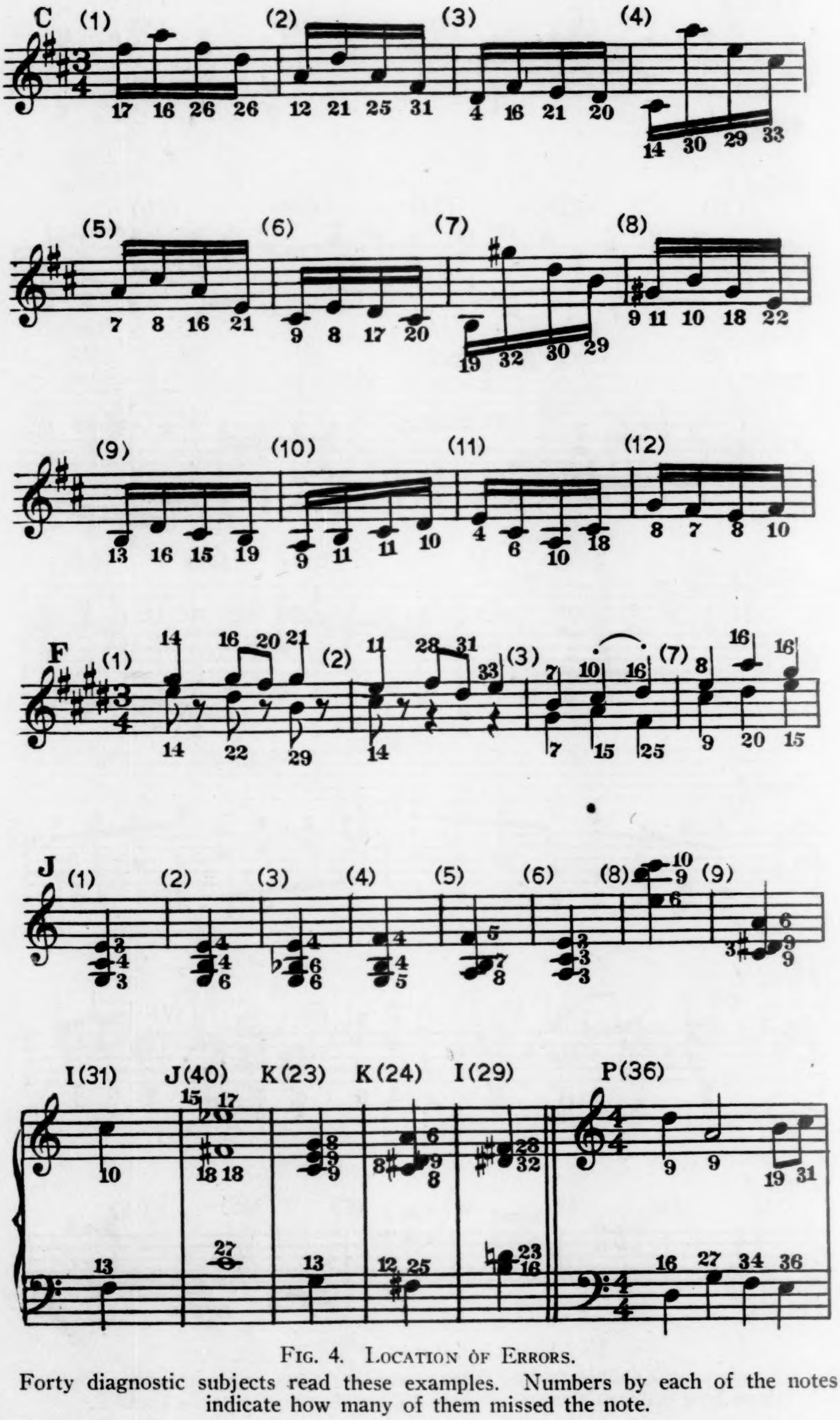


Fig. 4. Location of Errors.

Forty diagnostic subjects read these examples. Numbers by each of the notes indicate how many of them missed the note.

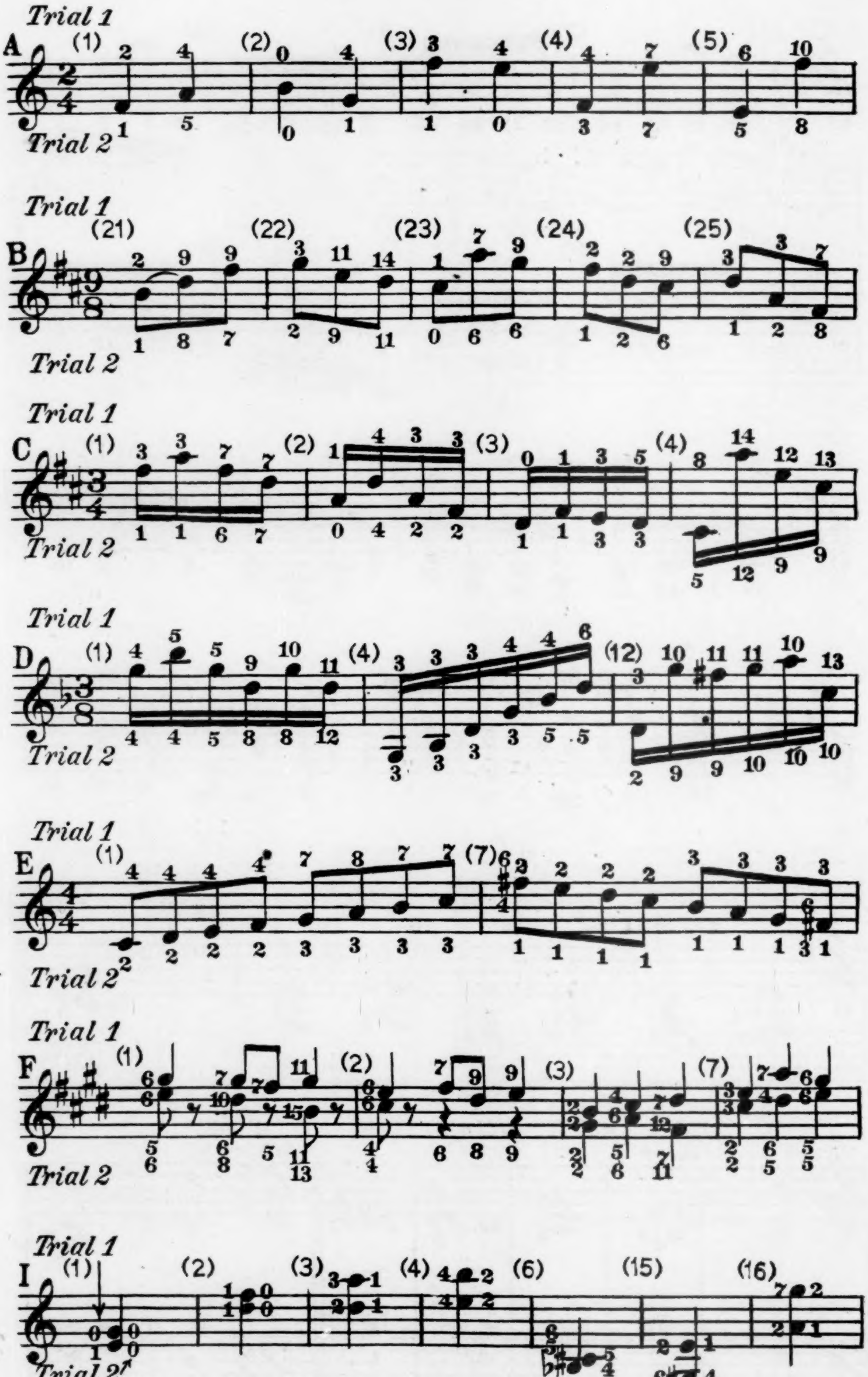


FIG. 5. ELIMINATION OF ERRORS.
Twenty-four practice subjects read these examples on two trials or more. Figures show how many of them made errors on each note for trials 1 and 2.

FIG. 5—CONTINUED

TABLE 2

Subject A—Training B.M. in violin, 1 yr. piano. Reading C—, playing C.

Trial	Time	Melodic M	Melodic R	Polyphonic M	Polyphonic R	Harmonic M	Harmonic R	Type of Reader
1.	92%	2.9	1-6	3.0	1-4	2.9	2-5	Part
2.	95%	3.1	1-6	3.0	1-4	3.1	2-5	Part
3.	96%	3.3	1-7	3.4	1-6	3.0	2-5	Pattern
4.	94%	3.5	1-6	3.3	1-6	3.3	2-5	Pattern
5.	97%	3.5	1-7	3.4	1-6	3.3	2-5	Pattern
6.	96%	3.5	1-6	3.2	1-6	3.4	2-6	Pattern

Final test of reading B—. Increased speed and accuracy.

Subject B—Training 10 yrs. piano (not with a good teacher, not recent). Reading E, playing D—.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	84%	2.1	1-4	2.4	1-5	2.4	1-4	Part
2.	88%	2.3	1-4			2.5	1-4	Part

Final test of reading D—. Insufficient speed, the same inaccuracy.

Subject C—Training B.M. in voice, 3 yrs. piano. Reading C—, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	89%	2.6	1-4	2.4	1-5	2.4	1-4	Part
2.	87%	2.8	1-6	2.4	1-5	2.3	1-4	Part
3.	87%	2.2	1-6	2.3	1-5	2.3	1-4	Part
4.	90%	2.5	1-6	2.4	1-5	2.5	1-4	Part

Final test of reading C. Very little if any improvement.

Subject D—Training B.M. in voice, 7 yrs. piano. Reading C, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	94%	2.7	1-6	3.0	1-6	3.4	1-6	Pattern
2.	96%	3.2	1-6	3.2	1-7	3.5	1-5	Pattern
3.	96%	3.8	1-8	3.6	1-7	3.8	1-6	Pattern
4.	96%	3.6	1-8	3.5	1-7	3.8	2-6	Pattern
5.	97%	4.1	1-8	3.9	1-7	3.7	2-6	Pattern

Final test of reading B—. Increase in accuracy chiefly.

Subject E—Training 9 yrs. piano. Reading C, playing B.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	95%	3.8	1-8	3.2	1-5	3.8	1-6	Pattern
2.	96%	4.1	1-8	4.0	1-7	3.9	1-7	Pattern
3.	96%	4.0	1-8	3.8	1-7	3.9	2-7	Pattern
4.	95%	4.0	1-8	4.1	1-7	3.9	1-7	Pattern

Final test of reading B—. Increase in speed and accuracy.

Subject F—Training 9 yrs. piano. Reading C—, playing B.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	91%	2.6	1-6	2.1	1-4	3.0	1-4	Part
2.	89%	2.9	1-8	2.4	1-6	3.1	1-4	Part
3.	92%	2.9	1-8	2.3	1-6	3.3	2-6	Pattern
4.	91%	3.2	1-8	2.6	1-6	3.5	2-6	Pattern
5.	92%	3.3	1-8	2.9	1-7	3.6	2-6	Pattern

Final test of reading B—. Greater speed. Accuracy slightly better, but still quite imperfect.

TABLE 2—CONTINUED

Subject G—Training 8 yrs. piano, considerable theory. Reading D, playing C+ (slow reaction time but accurate).

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	92%	3.0	1–6	3.5	1–7	3.2	1–5	Pattern
2.	93%	3.4	1–8	3.6	1–7	3.7	1–6	Pattern
3.	95%	3.6	1–8	3.8	1–7	3.8	2–6	Pattern
4.	95%	3.6	1–8	3.7	1–7	3.8	2–6	Pattern
5.	97%	4.1	1–9	4.4	1–8	3.9	2–6	Pattern

Final test of reading B—. Increased speed and accuracy.

Subject H—Training 7 yrs. piano, considerable theory. Reading D, playing B (very slow and inaccurate).

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	82%	2.4	1–6	2.3	1–4	2.5	1–4	Part
2.	86%	2.6	1–6	2.4	1–4	2.6	1–4	Part
3.	89%	2.8	1–8	2.4	1–4	2.8	1–5	Part
4.	87%	2.9	1–8	2.5	1–5	2.9	2–5	Part
5.	85%	3.0	1–8	2.7	1–6	3.1	2–5	Pattern

Final test of reading D+. Still very slow but more accurate.

Subject I—Training 7 yrs. piano, some violin, almost no theory. Reading C, playing B.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	94%	3.5	1–6	3.6	1–6	3.6	2–6	Pattern
2.	96%	3.6	1–6	3.8	1–6	3.6	2–6	Pattern
3.	96%	3.8	1–6	3.9	1–6	3.7	2–6	Pattern
4.	96%	4.0	1–8	4.0	1–7	3.9	2–7	Pattern
5.	97%	4.3	1–8	4.2	1–7	3.9	2–7	Pattern

Final test of reading B. Increase in speed and accuracy.

Subject J—Training 7 yrs. piano. Reading C—, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	87%	2.7	1–6	2.8	1–5	2.8	1–4	Part
2.	94%	2.9	1–8	3.0	1–5	2.9	1–4	Part
3.	93%	3.2	1–8	3.0	1–5	3.1	2–5	Pattern
4.	94%	3.3	1–8	3.1	1–5	3.1	1–5	Pattern

Final test of reading C. Increased speed.

Subject K—Training 6 yrs. piano. Reading D, playing B.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	94%	1.9	1–4	2.0	1–4	2.0	1–4	Part
2.	94%	2.2	1–8	2.2	1–5	2.7	1–4	Part
3.	94%	2.3	1–8	2.4	1–5	2.8	1–4	Part
4.	93%	2.5	1–6	2.5	1–5	3.1	1–4	Pattern

Final test of reading C. Speed increased, but inaccurate.

Subject L—Training 6 yrs. piano. Playing B, reading B—.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	95%	3.2	1–8	3.3	1–7	3.7	1–5	Pattern
2.	92%	3.7	1–8	3.6	1–7	3.8	1–6	Pattern
3.	94%	4.0	1–8	4.0	1–7	3.9	1–6	Pattern
4.	95%	4.3	1–8	4.1	1–7	4.0	2–7	Pattern
5.	94%	4.4	1–8	4.2	1–7	4.0	2–7	Pattern

Final test of reading B+. Practically flawless except time.

TABLE 2—CONTINUED

Subject M—Training 6 yrs. piano. Reading D, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	92%	2.0	1-4	2.1	1-4	2.3	1-4	Part
2.	91%	2.2	1-6	2.2	1-5	3.6	1-4	Part

Final test of reading D+. Fast and careless.

Subject N—Training 6 yrs. piano. Reading D—, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	85%	2.7	1-4	2.5	1-5	2.6	1-4	Part
2.	88%	3.0	1-6	2.7	1-5	2.8	1-4	Part
3.	92%	3.1	1-6	3.0	1-6	2.9	2-5	Pattern
4.	94%	3.1	1-6	3.0	1-5	3.0	1-5	Pattern
5.	92%	3.3	1-8	3.1	1-5	3.3	2-5	Pattern

Final test of reading C. More accurate, but many hesitations. Subject has some technical difficulty in playing.

Subject O—Training 5 yrs. piano. Reading D, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	89%	1.9	1-4	1.8	1-4	2.4	1-4	Part
2.	86%	2.4	1-5	2.0	1-4	2.5	1-4	Part
3.	89%	2.5	1-6	2.1	1-4	2.5	1-4	Part
4.	87%	2.4	1-6	2.0	1-4	2.7	1-4	Part
5.	89%	2.6	1-6	2.1	1-4	2.6	1-4	Part

Final test of reading C—. Increased speed, low accuracy.

Subject P—Training 5 yrs. piano. Reading D—, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	90%	2.8	1-4	2.5	1-4	2.9	1-5	Part
2.	91%	3.0	1-6	2.8	1-5	2.9	1-5	Part
3.	94%	3.3	1-6	3.1	1-6	3.2	2-5	Pattern
4.	93%	3.2	1-6	3.1	1-6	3.3	1-5	Pattern
5.	96%	3.4	1-8	3.3	1-6	3.6	1-7	Pattern

Final test of reading C. Accurate, but many hesitations.

Subject Q—Training 4 yrs. (Age 13 yrs., 8th grade.) Reading F, playing C. Reads one note at a time only.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	65%	1.5	1-4	1.5	1-3	2.0	1-4	Part
2.	75%	2.0	1-4	1.7	1-4	2.1	1-4	Part
3.	90%	2.1	1-5	1.6	1-4	2.1	1-4	Part
4.	86%	2.2	1-5	1.8	1-4	2.4	1-4	Part

Final test of reading D—. Still very slow and stumbling.

Subject R—Training 3½ yrs. piano. Age 14 yrs., 10th grade. Reading C—, playing B+. Considered a child prodigy.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	90%	2.9	1-8	2.6	1-5	2.5	1-4	Part
2.	92%	3.1	1-8	2.9	1-6	3.0	1-6	Pattern
3.	94%	3.1	1-8	3.1	1-6	2.9	1-5	Pattern
4.	94%	3.2	1-8	3.1	1-6	3.0	1-5	Pattern
5.	96%	3.6	1-8	3.5	1-7	3.5	2-6	Pattern

Final test of reading B—. Increased speed and accuracy.

TABLE 2—Continued

Subject S—Training 3 yrs. Reading C, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	93%	3.1	1-8	3.0	1-5	3.1	1-4	Pattern
2.	91%	3.5	1-8	3.3	1-6	3.5	1-5	Pattern
3.	93%	3.7	1-8	3.4	1-6	3.8	1-6	Pattern
4.	93%	3.4	1-8	3.4	1-6	3.8	1-6	Pattern
5.	95%	4.1	1-8	3.8	1-7	4.0	1-7	Pattern

Final test of reading B. Accuracy increased; speed slow.

Subject T—Training 3 yrs. piano. Reading C. playing B—.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	80%	2.4	1-6	2.3	1-5	2.5	1-4	Part
2.	87%	2.8	1-6	2.5	1-5	2.8	1-4	Part
3.	91%	2.9	1-6	2.9	1-5	3.1	1-4	Part

Final test of reading C+. Technique inadequate for speed.

Subject U—Training 2 yrs. piano, 2 yrs. voice. Reading D—, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	86%	2.1	1-4	2.0	1-4	2.1	1-4	Part
2.	88%	2.2	1-4	2.0	1-5	2.2	1-4	Part
3.	90%	2.2	1-6	2.3	1-5	2.2	1-4	Part
4.	93%	2.5	1-5	2.4	1-7	2.4	1-4	Part
5.	93%	2.8	1-6	2.6	1-6	2.7	1-5	Part

Final test of reading C—. Accuracy much greater, still slow.

Subject V—Training 2 yrs. piano. Reading D, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	89%	3.0	1-8	2.8	1-4	2.9	1-4	Part
2.	92%	3.2	1-8	3.1	1-5	2.9	1-4	Pattern
3.	90%	3.5	1-8	3.2	1-5	3.1	1-4	Pattern

Final test of reading C. Increased speed, still inaccurate.

Subject W—Training 2 yrs. piano. Reading C, playing C.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	94%	3.2	1-8	2.1	1-5	2.5	1-4	Part
2.	95%	3.5	1-8	2.1	1-4	2.6	1-4	Part
3.	95%	3.6	1-8	2.4	1-5	2.8	1-5	Pattern
4.	95%	3.7	1-8	2.4	1-5	2.7	1-4	Pattern

Final test of reading B—. Remarkable speed for two years training, but accuracy low.

Subject X—Training 1½ yrs. piano. Reading E, playing C—.

Trial	Time	Melodic		Polyphonic		Harmonic		Type of Reader
1.	70%	1.3	1-3			1.6	1-4	Part
2.	75%	1.5	1-4	1.1	1-3	1.7	1-4	Part
3.	79%	1.8	1-4	1.1	1-3	1.6	1-4	Part

Final test of reading D—. Reads one note at a time.

TABLE 2—Continued

Comparison of Group Mean Scores by the Critical Ratio Technique, All 24 Subjects

Time	Mean	S.D. of Difference	D./S.D.	Chances in 100 of a Difference
Trial 1.	88			
Trial 2.	89.8	4.18	.43	66
Melodic				
Trial 1.	2.6			
Trial 2.	2.88	.17	1.64	95
Polyphonic				
Trial 1.	2.49	.192	.99	84
Trial 2.	2.68			
Harmonic				
Trial 1.	2.71			
Trial 2.	2.89	.154	1.17	87
Same for 13 Subjects Who Practiced Five Trials				
Time				
Trial 1.	90.23			
Trial 5.	93.87	1.42	2.5	99
Melodic				
Trial 1.	2.75			
Trial 5.	3.57	.192	4.3	100
Polyphonic				
Trial 1.	2.707			
Trial 5.	3.39	.24	2.8	100
Harmonic				
Trial 1.	2.92			
Trial 5.	3.48	.148	3.8	100

notes not really seen clearly. How much guessing was done cannot be exactly determined, because unfamiliar groupings of notes in which guessing was impossible were not seen as wholes at all. They were deciphered, not read.

Experience in playing with a jazz orchestra resulted, in three cases, in rapid, but careless reading. These three individuals reacted quickly, but they did more surmising than perceiving. They failed miserably on melodic patterns and polyphonic music, but they found chords a little easier. They are listed in the table as J, M, and T. Subject J cured his careless tendencies, which were less in degree than those of the other two, who were careless throughout. Subjects J and T plunged ahead fast with little inhibition, and J especially took in large patterns at a glance, but both made numerous errors in details, especially in the melodic groups.

A few selected examples will illustrate how improvement was

recognized. Subject A, during her first trial, was able to grasp only one or two notes of a melodic pattern, usually at the beginning of the group. She showed by the errors that she made that she had no idea, as a rule, what the shape of the rest of the pattern was. She guessed at one or two notes beyond those she saw clearly, according to her report, and was right slightly more often than she was wrong. She reported a few times that she saw an entire group of six notes as a whole, but in every case the notes followed a scale or easy broken chord pattern up or down with no irregularities. She read polyphonic examples in the same manner. Chords for one hand appeared to her as wholes if very simple. Two-hand chords were rarely possible for her, and she reported that she guessed at one hand or the other, and did not see the entire chord clearly. Ultra-modern material did not register. One or two notes were all that she could read correctly, and no guessing was possible. The material in the comparison series gave further evidence for all of the above statements.

On trial 2 the errors were surprisingly similar to those on trial 1. Approximately 75 per cent of the cards that were wrong the first time were also wrong the second time they were read, and the mistakes were in the majority of these cases identical. There were a few cards incorrect on the second trial that were correct on the first, and a somewhat greater number of cases in which the reverse was true. A reported that she could see whole groups more frequently on the second trial than she did before, but that these patterns were so vague to her that she was never sure whether she got any of the notes right or not. In six such instances she played the entire pattern too high by one line, and in two cases too low by the same amount, but she grasped the relative position of the notes perfectly. She made only two responses similar to the above on the first trial.

On trial 3 for all types of material, A became a pattern reader predominantly, though she still resorted to analysis far too often. She reported that the patterns had ceased to be so vague to her, and she seldom misplaced them on the staff. She increased in both speed and accuracy on trial 4, but trials 5 and 6 showed no material gain except in time. Her final test without the appa-

ratus demonstrated performance which would not have been possible by her former method of reading. She was conscious that she was perceiving patterns as she went along, and that she was anticipating what she was about to play before playing it. The absence of any anticipation had caused stumbling in her reading previous to the practice. After the experiment, little trace of the stumbling was left.

Subject B, having neglected music for a number of years, gave unsatisfactory results, but managed to improve slightly.

Subject G was predominantly a pattern reader from the start, but she read slowly with fairly high accuracy. Her improvement consisted in learning to react more quickly, take in larger patterns at a glance, and read accidentals as a part of each pattern. She had the peculiar habit mentioned before of reading notes first and adding sharps, flats, or naturals afterwards, taking her time to the latter. Throughout the first three trials, a marked delay in reaction occurred every time a sharp, flat, or natural occurred. More than one accidental persistently resulted in loss of the impression of what was read before a response could be made. When no accidentals were present, the reaction time was short by the second trial, and the response was accurate, even if the pattern included seven or eight notes if these were not too irregular in arrangement. The subject could read two-hand chords far better than the average provided not more than one accidental were present. On trial 5 for all types of material, G showed marked improvement in her peculiar disability, and reported that she had become able to grasp sharps, flats, and naturals along with the notes more often than before. The final test of her reading showed that she was still too slow, but had gained some in speed, and considerably overcame her hesitant manner of reading before the experiment, which may have been caused partly by nervousness.

Subject H, who has been mentioned before, was considered gifted, had a phenomenal memory, and played with facility a broad repertoire. Several outstanding teachers have attempted to correct his pronounced reading disability. With his own persistent coöperation, all efforts failed completely. His improve-

ment was gradual. At first he used one or two notes as cues, and typically spent over two seconds figuring out his response before he could begin to play. Any rhythmic patterns were likely to give him trouble, and he hesitated frequently, seldom keeping a steady tempo. Until the last trial, he could not read chords for both hands with any certainty at a glance. He persistently failed with unusual chords throughout the practice, though he has memorized a considerable amount of contemporary music from several composers. Near the end H averaged slightly more than one second in his reaction time, and he reported that he saw the groups of notes as wholes on a little over half of the cards. The location and nature of his errors confirmed this report. However, his inhibitions and slow, careful habits seemed to persist to some extent in his sight-reading after the experiment.

Interesting as compared with H is subject I, whose reading was superior to that of H in every respect even though her training was inferior. She had studied very little theory, put no extraordinary effort into improving her sight-reading, memorized comparatively little, and played almost no music in ultra-modern idioms. She depended upon the printed page nearly always when she played. This last fact may account partly for her efficiency in reading, but it does not explain how she can grasp some unusual and irregular groupings which are probably foreign to her experience, while H, who must have played many such patterns, persistently failed to read them. The improvement of I consisted in increasing the span of perception, and her progress was steady.

The best reader in the entire practice group was the gifted 17-year-old pianist L. She grasped patterns of three or four notes with ease from the start, and later took in a six-note span with a surprising accuracy approaching that of the best professional musicians in the diagnostic group. She shortened her reaction time during practice, and became a rapid as well as accurate reader. She made approximately three-fourths of her errors persistently on three or more of the five trials. Improvement seemed next to impossible on all material that was spread out over a wide area. Series M and the first part of Series B

gave her the most trouble for this reason. The last 20 cards of Series A failed to improve much for the same reason. L read two-hand chords with ease unless they covered a vertical distance of 1¾ inches or more. She read nearly all of Series L, consisting of unusual chords, with less difficulty than any other subject on all trials, and nearly equaled two experienced professionals on her last trial. She, like subject I, could give no clue introspectively as to how she perceived these unusual groupings, which she often could not name as altered dominant sevenths, ninths, or whatever they were. She claimed that they had meaning for her, however, even if she could not identify them easily as to harmonic structure.

Some of the subjects, as may be seen from Table 2, made steady advancement, while others were inclined to reach a plateau or to backslide occasionally during practice. Interest was uniformly good throughout, except as it was affected by moods or health in rare instances. The writer would infer that on the last trial the subjects were apparently stimulated to greater effort in a few cases by the feeling that they must make the best of their last chance to improve. A few records indicate such a motivation, which ended a plateau with a sudden jump to much higher efficiency. After the practice several subjects were amazed at their own increase in speed and accuracy of sight-reading. One stated that he could now read so fast that his fingers could not keep up with his eyes. Others expressed their reaction to their gains in efficiency with words to the same effect.

V. Conclusions

Within limits determined by the kinds of music that were used, the results of this investigation prove conclusively that in the group of 74 subjects tested, which we may assume to be typical of professional, student, and amateur musicians, the ability to grasp a musical pattern of three or more notes at a glance when necessary is absolutely essential to rapid and accurate reading. No exceptions to this rule were found, and Jacobsen's results tend to support this conclusion.

Readers who typically see one or two notes which serve as sensory cues fill in the remainder subjectively. They read only part of the pattern and guess the rest of it. They may therefore be classified as part readers, and they are, without exception, less efficient than pattern readers, who use the form of the whole as a cue to recognition for most of the material that they read. Other things being equal, the larger the patterns grasped, the faster is the maximum speed of accurate reading.

There seems to be a very low positive correlation between years of musical training and span of perception in the tables for all types of material, but a slightly higher correlation in the case of polyphonic music than for the other two classes. Since years of study make a poor index to achievement, and no more accurate measure of accomplishment was applied to all subjects in this research, the writer feels that the exact coefficients of correlation would mean little with the data for their calculation. Therefore we may omit detailed consideration of this matter, and proceed from these generalizations to more important specific conclusions.

1. *Musical Patterns that Are the Equivalent of Words.*—The influence of background is evident in the successes and failures of individual subjects on different cards. Each subject had a musical vocabulary, if that term may be used in such a connection, which may be said to correspond to his English vocabulary in that configurations frequently encountered in past experience evoked immediate recognition just as a familiar word would have done. A relatively unfamiliar word, on the other hand, must usually be subjected to analysis before its meaning is recognized, and an unusual musical figure must be treated in the same manner. We refer, for proof of this, to the preceding chapter in which was mentioned the fact that several subjects reported that for some cards the meaning was not immediate. It was nonsense to them until they subjected it to analysis, and learned how it sounded.

Several degrees of difficulty of perception should be distinguished in order to show what constitutes complexity in printed music. A few general considerations will be given first, then more specific conclusions in regard to this matter. The duration

of the notes plays the major rôle in determining what will be the perceptual unit. In case of half notes played at a very slow tempo, one note constitutes such a unit, according to the results with Series B, cards 1 to 20. The time signature determines the groupings to some extent. The distribution of notes on the staff is important. In Series A, for example, card 21 (see Fig. 2), the average reader failed to see the F clearly. He saw only the first two notes as a unit, and anticipated the third note out in the periphery as belonging to a separate unit to be shifted to later. With eighth notes crowded more closely as in Series B, card 21, the typical reader saw all three notes as a unit. Therefore number of notes alone does not determine difficulty. In Series E, card 1, for example, any observer who had played scales from the page to any extent recognized the scale on the card, and got the eight notes right if he did not misplace the entire scale. All that he was required to do in order to respond correctly was to use the form as a cue and be careful to observe where on the staff the scale started. In the case of D 12, for example, the pattern consisted of only six notes, but the arrangement was irregular and unusual, and the best pattern readers sometimes missed it.

By analysis of the errors on numerous examples with all subjects for several kinds of material, the experimenter was able to arrive at a number of interesting, more specific conclusions.

For melodic material, the following was the order from easiest to most difficult patterns in general for a typical reader: ascending stepwise progressions, descending stepwise progressions, ascending broken chord progressions following a simple major or minor triad, descending examples of the same, ascending chromatic progressions, descending chromatic progressions, alternations between two notes a third or more apart, broken seventh chords, broken altered chords, irregularly ascending groups skipping sometimes one step and sometimes two, descending examples of the same, and irregular groupings with some wide skips and one accidental.

The operation of two factors was clearly seen to control the difficulty of the material. One was the frequency with which

the pattern occurred in the musical experience of the observer. Another was the number of elements, including sharps, flats, and naturals, that composed the pattern. Chromatic passages or altered chords contained accidentals, which increased the difficulty. Otherwise frequency explains the order of difficulty except for the curious fact that ascending passages were read with greater success than descending passages, while both occurred about an equal number of times during the experiment. The difference was small, but reliable in each case. The writer believes that the slightly greater difficulty of playing descending passages on the piano experienced by performers not too highly skilled may have resulted in worse reading habits coming down than going up, but this is only a tentative explanation.

The six-note groups were more easily grasped than those of five or seven notes. Five were usually played as six, and seven as eight, or six. Here again the rarity of fives and sevens in music played by the observers operated as a confusing factor. Also the four-, six-, and eight-note gestalten are frequently capable of being subdivided into smaller groups. Six notes may appear as two groups of three or three groups of two, and eight as two groups of four or four groups of two, etc. This subdivision facilitates reading.

Rhythms involving quarter notes and eighths interspersed were more easily read than the same sort of patterns with eighths and sixteenths. In the latter case the distinction was difficult because the number of flags on each stem had to be observed. If bars connected the notes instead of flags being placed on each individual stem, the perception of rhythms was easier, because the bars unified the whole by binding the notes together. This Gestalt principle is often encountered in other categories of objects. Most printed music is patterned conveniently for reading of rhythmic groups, but exceptions can be found.

Other things being equal, most subjects had greater difficulty with the bass than with the treble staff, probably because the latter was relatively less familiar to them.

As might be expected, the greater the number of sharps or flats in the signature, the more confusing was the material, and

without exception sharps gave the subjects more difficulty than flats. The more sharps or flats the observer must keep in mind, the more complex is his task, even though he is not required to perceive more elements. Therefore a mental set for four sharps, for example, adds new elements in the mind of the observer which would not be there in the key of C major at all. Why flats should ordinarily be less difficult than sharps is not clear to the writer.

Parallel motion of two melodic lines was more easily perceived than contrary motion. Probably the relatively greater frequency of parallel motion in most music accounts for the fact that notes fit together into a whole better in this kind of structure. However, parallel motion at the interval of a fourth or a seventh caused great confusion. The former were most frequently misread as thirds, and the latter as octaves. The readers tended to conform to the pattern which they had frequently played which closely resembled what they saw. Parallel fourths or sevenths seemed illogical to them.

Though intricate rhythms could often be read in one melodic line, any simple lack of coincidence of the time values in two-part polyphonic selections gave trouble. There were certain notes in one part or the other which seemed to stand outside of the pattern. They seemed to be excluded, as it were, from the unit being read. Therefore perception at a glance of the time relations of two simultaneous melodic lines were almost impossible except for the very best readers. The conclusion that certain notes were excluded from the Gestalt at times is further substantiated by the fact that these notes were nearly always misplaced, while the other notes in the same group were correct.

More than two parts in counterpoint did not seem to be read in patterns consisting of one or more notes from each of the melodic lines unless three or four notes fell in a vertical arrangement as in a chord. If this vertical arrangement did occur, the chord was read as a whole, but if the fitting together of the different tone values was at all intricate, the best readers failed to take in all parts at a glance. Probably frequent shifts from one part to another were always necessary, as Jacobsen's results indi-

cated, but how rapid and accurate reading is thus possible with such material remains a mystery to some extent. The reason for the failure to grasp all parts at a glance is clear since the notes are separate, distinct units, unconnected usually by stems or any other lines that would tend to unify them into patterns.

The fact that notes on lines appeared to be different entirely from those in spaces would indicate that the staff was part of the figure and not ground in the usual Gestalt sense of the term.

With two-note chord material for one hand, the intervals in order from easiest to most difficult to read were as follows: thirds, fifths, octaves, seconds, fourths, sixths, and sevenths. This conclusion is only tentative, however, since not enough examples of each interval were used to eliminate other factors, such as position on the staff, etc. Very little experimental evidence can be given to support any explanation of this order of difficulty, which may have been partly a matter of mere chance. However, the writer believes that when both notes were on lines or both in spaces as in the case of thirds and fifths, the mere likeness of the elements facilitated recognition. Octaves came next probably because of their frequency in music. Sixths were wide apart, had unlike elements—one note on a line and the other in a space—and were less frequent than octaves. Sevenths had like elements, but spread far apart. They were often mistaken for octaves, which are more frequent in occurrence.

Triads were most easily read in their fundamental positions, in which all three notes are on lines or all in spaces, and even distances apart. Neither of these conditions holds in other positions, in which one note is always removed a fourth from the other two. More complex chords followed the same laws. Some elements on lines and others in spaces added to the difficulty of perception. Distribution at equal intervals apart formed a better Gestalt than uneven distribution. As said before, seventh, ninth, and altered chords were more difficult than triads because there were more elements to be seen, and besides this the average reader in the group was not much experienced with seventh, ninth, or altered chords. There was some difference also in the unity of the pattern depending upon whether the elements were

connected by one stem or separated into two parts for two hands. The more unusual chords gave more trouble for this reason than did the triads.

Reading at a glance, subjects were able to grasp a few chords written on one staff for the right hand easily, while they experienced difficulty with the same combinations written in such a way that two notes were on one treble staff, and one, or possibly two, on another treble staff. The larger field that must be covered when reading for two hands at once may account for this. However, another factor seems to have entered into this situation, and that was the strong tendency that nearly all subjects had to think of the left-hand part as being on the bass staff. The uncommon use of treble staff for both hands was rare in their experience, and when it had occurred, slow deciphering was the result. A few examples proved that the same left-hand part could be read with greater ease in the bass than it could in the treble, even though the notes were identical.

The vast majority of the observers could respond most easily by playing on the keyboard. Therefore any difficulty with the motor act in the response was due to insufficient training, and only two cases of such a difficulty were discovered. Both were slight. If any difficulty was experienced by other observers in the transfer of the visual perception into a response, the trouble was obscure so far as the observers were concerned. The performance of observers on comparison cards, in so far as an accurate measure of it was possible, matched so closely the results at the piano that the difficulty in transfer to the keyboard seemed of minor importance as a rule. However, other phases of the reading process involving imagery and anticipation of action are yet to be worked out.

The problem of the relative importance of optical complexity and musical complexity offers a number of rather profound aspects. First in importance among these aspects is the question of which of these two ways of thinking of the musical pattern was dominant in the mind of the observer when he perceived the pattern. A question of this nature is usually difficult, if not impossible for the observer to answer subjectively. Some indi-

viduals claim that the visual pattern is what they look for in music, and that just how the visual meaning is transferred over into action is a mystery to them. Others state that they probably hear, at least vaguely, how the music sounds before they play it. This step in the transfer to the keyboard is impossible for some individuals, according to their reports, and each visual stimulus of a note means to them that they must press the appropriate key and that is all. They frequently express surprise at the unexpected sound that results.

Therefore the manner of transfer of visual perceptions to the keyboard remains a separate problem. This research goes far enough to suggest that wide individual differences in imagery determine whether optical or musical complexity of a Gestalt has the more confusing effect upon a subject. In the comparison series of this research were given two examples of stimuli in pairs in which the first of the pair was like the second in notes, but different only in visual appearance. In one case the only difference was that part of the stems were turned upward in one pattern and downward in the other. In the other case an arpeggio was printed with the notes distributed evenly for one stimulus, and with the same notes with one irregularity in their distribution on the other side. Those observers who appeared to be dominantly visual minded discovered the exact nature of the difference in each case, while those who were most concerned with the sound of the music invariably thought that different notes were in the two stimuli in both examples. They were unable to state what notes were changed, however, except in two or three cases in which the observers were sure that certain notes had been changed, and named the notes.

Another example is of value in illustrating the confusing effect of successive, optically dissimilar, but musically identical patterns. In Series I there was a two-note chord written A flat E flat which was followed by the chord G sharp E flat. Obviously both sound alike on the piano, and are played by pressing the same keys, yet confusion and delay occurred when the latter was seen because the combination of a sharp and a flat in one chord is rare in the sort of music most of them had played.

The inability to discriminate at a glance between an ascending scale passage and an ascending arpeggio may be regarded as a serious symptom of inaccuracy in reading. Such faulty observation was common among the less efficient readers until practice eliminated it in most cases. Similar inaccuracies were found in the reading of small irregularities in an otherwise uniformly ascending or descending progression. These also improved with practice.

The conclusion which may be tentatively proposed after careful analysis of these and other examples is that the readers who were habitually inclined to read one or two notes at a time, when forced to grasp larger units of material, saw only vague forms at first. They derived from their rather uncertain impression an idea of the direction of motion, but not the distance over the staff that the progression went, or even the intervals between the notes. Hence they often played too many notes or too few, totally disregarding the time. This often resulted in no improvement so far as figures in the table were concerned, but it should be regarded as the first step forward in learning to read by groups instead of single notes. There were several subjects who had advanced beyond this stage before starting the practice in this experiment, and these individuals saw the intervals between the notes more clearly and as a rule played the correct number of notes. The increase in clearness of the vague impressions of patterns, which many received at first, was gradual. If the practice had been continued for a longer period, a greater accuracy could have been obtained, the experimenter believes, than was possible at the end of the limited time. The first improvement that was evident was the larger units that the observer grasped, usually near the beginning of the second trial. After that the accuracy increased slowly. During the later trials the complexity of patterns that could be seen as wholes increased only a little, but more details were accurately observed than on earlier trials.

The fact that entirely new habits were set up as a result of practice with the tachistoscope is evident from the results of most

subjects. Why a few failed has already been explained, but why some started with correct methods of reading before the practice is unknown, since others under the same instructors in music had set up wrong habits. The methods of the former individuals were an accident. They were unconsciously picked up in spite of the over-analytical teachers.

The final test of reading with some new cards and some old cards revealed that the practice effects were general. As stated earlier, the subjects were as a rule unable to distinguish between old material that they had practiced and new material that they had not seen. The span of perception was about the same for both, though not enough new material of each type was used to furnish accurate figures for the table comparable with those for each of the practice trials. However, the final test showed that the new habits were generally applied effectively to any music to be read.

The close analogy between the reading of music and the reading of language has been clearly brought out in connection with the facts found. The good music reader must grasp units of four, six, or even eight notes just as he reads so many letters grouped together as a word. Therefore appropriately grouped notes form units that are in visual or auditory perception, or in meaning as related to their context, the equivalent of words in prose or poetry. They should then be read as such.

2. *Practical Application of This Research.*—The data obtained from the diagnostic and practice experiments suggests that a measure could be made of more individuals at each level of advancement in order to standardize a test of music-reading efficiency that could be applied by teachers. The same or similar material could be used with a much simpler means of exposure. The writer is attempting to construct a simple tachistoscope that may be mounted on the piano easily and used by teachers. He would propose that each teacher of piano make a record of the average span of perception of every pupil for different kinds of material. The types of errors made frequently by each should also be recorded. Then the resulting figures should be compared

with figures recorded in Table 1 and Table 2 of this investigation. The writer hopes that later some reliable norms will be available for comparison of the individual with a large number of others at his level of advancement in musical education.

If disabilities show up in the scores of any pupil, considerable time should be devoted to practice in reading cards in the tachistoscope until the pupil is forced to abandon unsuccessful methods of reading in favor of pattern reading. After considerable improvement is in evidence, a second test should be made to determine quantitatively the increase in complexity of patterns that can be used correctly. This will show the pupil his own improvement in clearly understandable form. The practice should not cease after this test. Much easy material should then be read at sight daily with the Gestalt principle in mind. If any sign of return to the old habits is observed, the teacher should practice the pupil with the tachistoscope again until backsliding has been checked. After several weeks' application of the new approach to continuous reading, another test should be made to make a quantitative measure of the improvement during that time.

Probably the tachistoscope measurement and practice is of little if any value until the pupil has automatized to a fair degree his response to single notes, even if he has not yet learned thoroughly signatures or accidentals. The teacher should see the destructive effect of continually calling the pupil's attention to the single note by scolding him whenever he happens to miss one. Corrections are necessary at times, but frequently the pupil should be urged to drive ahead rapidly with as few inhibitions as possible, even if he does make a few errors in the notes. In the end he will gain by acquiring skill in reading fast and accurately, while if he is constantly inhibited by fear of playing wrong notes, he will play the correct notes slowly and stumblingly.

The same procedure can be followed, but with different material, for the teaching of sight-singing. Educators are working in the right direction in this field already, but they have not gone far enough yet to have any accurate measure of the results that they are getting by pointing out tone groups to the children, or showing them flash cards. The groups that they are dealing

with are mostly phrases or logical divisions of the music, not perceptual units. The two are seldom identical. Further experiments with flash cards in the school room with controlled exposure time, may demonstrate that pattern reading can be taught along with sight-singing. However, reading the singing material on flash cards will probably prove insufficient for those who study piano or organ. Instrumental music requires quick grasp of rapid passages that could never be sung, and also of polyphonic selections and chords. The performer must learn to respond quickly upon his instrument, and to anticipate beforehand what he is to do.

Ordinarily musical material for the beginner is fairly well selected or composed so far as stages of difficulty in reading or playing are concerned. Little criticism of public school music material is to be made on this point. The groupings are printed in a manner that has been found by experience easiest to read, ordinarily. Only the method of teaching these materials should be changed in order to fall more in line with the Gestalt approach. The content rather than the arrangement of some of the music should be criticized, since frequently it is decidedly uninspired and uninspiring in nature.

The writer hopes that music reading may be made as enjoyable to everyone who can sing or play as novel reading is for most educated people. In order for it to be a joy instead of a pain, as it now is to many, this process must be made simpler, so that excessive strain is avoided, and pleasing results come as a reward for one's efforts. Training in the Gestalt approach to printed music should make this skill possible for many singers and players who would otherwise avoid reading whenever possible. No claim is to be made for such training as a remedy for all reading disabilities, but a cure for some inefficient habits has been proved possible in this research.

VI. Résumé

This investigation is a study of the complexity of musical patterns that can be perceived with one fixation of the eyes by

individuals with various degrees of musical training and experience, and of the effects of tachistoscope practice on the span of perception of these individuals for various kinds of musical material.

The tachistoscope mounted on a piano proved itself of value in measuring music reading efficiency for different kinds of material. The diagnostic experiment, with 50 subjects having various amounts of musical training from that of the professional to that of the beginner, proved conclusively that without exception efficient readers are able to grasp groups of three, four, or more notes at a glance. Slow, stumbling readers, on the other hand, are seldom able to grasp more than one or two notes at a time. On the basis of this distinction in fundamental reading habits, we may classify those who use the form of the entire group of notes as a cue to recognition as pattern readers, and those who see only one or two notes of a group and fill in the remainder subjectively as part readers. There is a low positive correlation between years of musical training and span of perception, or the amount of material that may be grasped as a unit at a short exposure. However, accomplished professional readers are able to read large patterns accurately at a glance even in the most intricate compositions.

Many excellent pianists and violinists are part readers, and their individual deficiencies proved quite interesting. Not all of them had the same disabilities, but in general they tended to spell out music or decipher it instead of really reading it. The inefficient habits that they had seemed to be rooted in their early training, and they had little or no insight concerning their difficulties. Several individuals were found who were not advanced in musical education, but were remarkably rapid and accurate readers, considering their background. They read mostly by patterns. They were totally unaware of the methods that they used as well as of the origin of their efficient habits.

Practice with the tachistoscope demonstrated that part readers may become pattern readers if their response to notes is sufficiently automatized. Approximately 20 to 30 hours of tachisto-

scope reading were required to make a substantial improvement in most subjects. Span of perception increased first and accuracy later in nearly all cases. The pronounced progress was evident to the subject in nearly every case, and in some instances the individual, who was formerly a stumbling reader, became able to read so fast that his fingers could not keep up with him. The experimenter's judgment as to the amount of improvement was in several cases confirmed by teachers who had the subjects under their instruction in piano.

With complete data available on every subject's errors with each card, the experimenter was able to classify his material as to the difficulty of various kinds of patterns. The results of a careful study of the kinds of errors that were persistent led to a Gestalt interpretation of music reading, which should be applied to learning in this field just as Gestalt already has proved valuable in teaching children to read language by large units. From the data we may conclude that a close analogy exists between the reading of words and the reading of music, for there are groups of notes which are definite perceptual units, and in this way the equivalent of words. Every musician has a musical vocabulary just as he has a vocabulary in English, French, or any other language. Any unfamiliar idioms must be deciphered until they are learned, but well known patterns should be recognized at a glance. Sometimes, however, wrong habits make immediate perception impossible.

This study has only attempted to classify music reading disabilities on the basis of errors made with varied patterns. There are numerous problems in this field which must be solved before the entire process of music reading is understood, but this investigation has stressed the importance of a neglected step in learning this skill of which most music teachers are ignorant, namely, the Gestalt approach to the process.

VII. Bibliography

1. Buswell, G. T., How perception operates in reading, from Robinson and Robinson, *Readings in General Psychology*, Univ. Chicago Press, 2d ed., 1929, pp. 295 ff.

2. Buswell, G. T., *Fundamental Reading Habits: A Study of Their Development,* Dept. of Educ., Univ. of Chicago, 1922.
3. Dearborn, W. F., The tachistoscope in diagnostic reading, *Psychol. Monog.,* 1936, **47**, No. 212, 1–19.
4. Dodge, Raymond, An experimental study of visual fixation, *Psychol. Monog.,* 1907, **8**, No. 35, page 95.
5. Glanville and Dallenbach, The range of attention, *Am. J. Psychol.,* 1929, **41**, 207–236.
6. Hofman, Joseph, *Piano Questions,* Doubleday, Page, 1909, page 117.
7. Jacobsen, Ole, An analysis of eye movements in reading music, Thesis, University of Chicago, 1926.
8. Mursell, James L., *Principles of Musical Education,* Macmillan, 1937, Ch. VIII, 141–168.
9. Ortmann, O., Elements of chord reading in musical notation, *J. Exper. Educ.,* 1934, **3**, 50–57.
10. Robinson, F. P., The rôle of eye movement habits in determining reading efficiency, *Proc. Iowa Acad. Sci.,* 1932, **39**, 231–235.
11. Robinson, F. P., The tachistoscope as a measure of reading perception, *Am. J. Psychol.,* 1934, **46**, 132–135.
12. Sisson, E. D., The rôle of habit in eye movements in reading, *Psychol. Record,* 1937, **1**, 157–168.
13. Stone, Clarence R., *Better Primary Reading,* Webster Publishing Co., 1936.

...a Concludes Volu... ...ains
...ndex for the Ent...

...No. 228
...1938

...logical M...raphs

Edited by
John F. Dash...
University of North ...

...xperimental ... to
...e Reading o...

By

Kenneth L. ...

...n Submitted in P... ...llment of
...rements for the D... ...tor of
...phy in the Univers... ...igan
1938

Published by
...ican Psychologica... ...ion, Inc.
...e Ohio State University, ...